



# Dynamic effects of price promotions: field evidence, consumer search, and supply-side implications

Andrés Elberg<sup>1</sup> · Pedro M. Gardete<sup>2</sup>  · Rosario Macera<sup>3</sup> · Carlos Noton<sup>4</sup>

Received: 3 November 2017 / Accepted: 27 August 2018 / Published online: 4 October 2018  
© Springer Science+Business Media, LLC, part of Springer Nature 2018

## Abstract

This paper investigates the dynamic effects of price promotions in a retail setting through the use of a large-scale field experiment varying the promotion depths of 170 products across 17 categories in 10 supermarkets of a major retailer in Chile. In the intervention phase of the experiment, treated customers were exposed to deep discounts (approximately 30%), whereas control customers were exposed to shallow discounts (approximately 10%). In the subsequent measurement phase, the promotion schedule held discount levels constant across groups. We find that treated customers were 22.4% more likely to buy promoted items than their control counterparts, despite facing the same promotional deals. Strikingly, the magnitude of the dynamic effects of price promotions (when promotional depths are equal across conditions) is 61% of the promotional effects induced by offering shallow vs. deep discounts during the intervention phase. The result is robust to other concurrent dynamic forces, including consumer stockpiling behavior and state dependence. We use the experimental variation and historical promotional activities to inform a demand-side model in which consumers search for deals, and a supply-side model in which firms compete for those consumers. We find that small manufacturers can benefit from heightened promotion sensitivity by using promotions to induce future consideration. However, when unit margins are high, heightened promotion sensitivity leads to fierce competition, making all firms worse off.

**Keywords** Price promotions · Dynamic effects · Consumer search · Bertrand supertraps · Field experiments

---

We thank Eva Ascarza, Bryan Bollinger, JP Dubé, Wes Hartmann, Avery Haviv, Przemyslaw Jeziorski, Carl Mela, Ilya Morozov, Harikesh Nair, Duncan Simester, Miguel Villas-Boas, and the review team for helpful comments, as well participants in research seminars at Pontificia Universidad Católica de Chile, Santa Clara University, UC Berkeley, Universidad Diego Portales, and at the Marketing Science Conference, QME Conference, SICS, the UCSD Rady Field Experimentation Conference, and the Workshop in Consumer Analytics. The outstanding research assistance by Claudio Palominos and Breno Vieira is gratefully acknowledged. Noton acknowledges financial support from the Institute for Research in Market Imperfections and Public Policy, ICM IS130002.

Extended author information available on the last page of the article.

## 1 Introduction

Consider two hypothetical retail customers, in all identical except in the promotions they face. For example, one might be offered deeper discounts for a set of goods than the other for a period of time. Even though we can expect that during the promotional period the first consumer will demand more goods than the second, less is known about how they will react to subsequent promotions and how the reaction will affect the profitability of firms competing for their business.

The likely direction of dynamic effects of promotions is not straightforward. On the one hand, one can imagine the scenario in which consumers become “addicted to deals.” As a result, a period of attractive promotional incentives may *dampen* consumers’ sensitivities toward subsequent promotions, and firms may be required to ratchet incentives up to maintain buying behaviors. This prediction is in line, for example, with some of the results of the rational addiction literature (e.g., Becker and Murphy (1988) and Becker et al. (1991)). On the other hand, consumers might display “heightened promotion sensitivities.” In this case, promotional incentives *increase* consumers’ future promotion sensitivities. As a result, firms may be required to offer lower promotional incentives to maintain customers’ purchasing activities. The habit-formation literature has proposed these types of dynamic effects (e.g., Pollak (1970) and Spinnewyn (1981)).

We investigate the dynamic effects of price promotions by use of a large-scale field experiment. In collaboration with a major retail chain in Chile, we exogenously varied the prices of 170 products across 17 categories in 10 supermarkets to study how the depth of present promotions affects consumers’ sensitivities to subsequent promotion activities. The experiment is organized in two halves. In the first half (the intervention phase), a set of products sold in five *treated* stores received “deep” discounts (approximately 30%) according to a promotion schedule. Simultaneously, the same products were sold under the same promotion schedule at “shallow” discounts (approximately 10%) in five similar *control* stores. In the second half (the measurement phase), the same set of products was promoted with identical 10% discounts across all stores. Our focus is on systematic consumer behavior differences occurring during the second half of the experiment, as a function of the conditions assigned during the first half.<sup>1</sup>

We also take advantage of the exogenous variation generated by the experiment to inform and investigate a model in which consumers search for deals. Together with historical price data, used to inform consumer beliefs, the model allows us to consider the implications of consumer search for competing manufacturers. We develop a competitive supply-side model in which manufacturers compete over customers through promotional activities.

Our experiment reveals that consumers exposed to deep promotions exhibit subsequent heightened promotion sensitivity. In particular, treated customers were 22.4%

---

<sup>1</sup>The marketing literature distinguishes between several price promotion instruments, including temporary price reductions (TPRs), coupons, promotion packs, rebates, among others. Our focus is on TPRs – the most frequently used type of price promotion (Gedenk et al. 2010).

more likely to buy promoted items during the second half of the experiment than their control counterparts, despite being exposed to the same promotions at the time. Strikingly, the dynamic effect constitutes 61% of the static effect induced by the different experimental conditions in the intervention phase. Exposure to deep promotions during the first half of the experiment also increased the proportion of promoted goods in treated consumers' baskets in 4.7% during the measurement phase.

Our findings motivate the interpretation that offering attractive promotions today can act as an invitation to search for deals to a greater extent tomorrow. In particular, we find that promotion depths are positively correlated over time, which implies that finding a deep discount on a given week is a positive indicator of the likelihood of a deep discount occurring in the following one. Under beliefs consistent with these data, our model predicts that consumers in treated stores should search and buy a higher share of promoted goods than their control counterparts. Regarding the supply-side, the counterfactual analysis reveals that small firms, whose products are the least searched, have a higher incentive to provide deep promotions to generate future consumer consideration. However, these firms are also the most penalized once competitors also find it valuable to offer deep promotions. Also, we find that all firms can become worse off when consumers exhibit heightened promotion sensitivity, due to the resulting increase in competition intensity.

This paper contributes to the available literature by specifically studying the effects of current promotions on the sensitivity to future promotions. Previous literature has mostly focused on other dimensions of consumer response, such as the impact of promotional activities on future purchasing behavior. Approaching this research question, predominantly through the analysis of observational data, early work by Scott (1976), Shoemaker and Shoaf (1977), and Dodson et al. (1978), Jones and Zufryden (1980) and Guadagni and Little (1983), as well as later work by Gedenk and Neslin (1999), report that repeat purchase probabilities tend to be lower when following a promotion purchase compared to repeat purchase probabilities following a non-promotion purchase. Neslin and Shoemaker (1989) and Davis et al. (1992) and Ehrenberg et al. (1994) on the other hand, offer some skeptical views on the negative association between purchases during promotions and subsequent purchases.<sup>2</sup>

An exception to the use of observational data to study the effects of promotions on future purchases is the experimental evidence in Anderson and Simester (2004). In the context of mail-order catalogs selling durable goods, the authors randomize promotions at the individual level, and find that new customers who are offered deeper discounts purchase more in subsequent orders relative to the control group. In contrast, established customers react in the opposite way, by reducing their subsequent purchases. The analysis by Anderson and Simester (2004) focuses only on customers who selected into the sample by purchasing an item during the intervention period.

---

<sup>2</sup>Other dimensions of the effects of promotions have also been studied. For instance, Kopalle et al. (1999), using a descriptive dynamic sales model, find that the high- and low-share brands tend to over and under-promote, respectively. Erdem et al. (2008) develop a structural model to show that, since price is an important signal of quality, frequent price cuts can have an adverse effect on brand equity, while Chan et al. (2008) use a structural model to show that future reactions to current price promotions depend on whether consumers are brand loyal or not. See Neslin and Van Heerde (2009) for an extensive review of promotion dynamics.

The authors find suggestive evidence of selection, which they control for through the use of historical controls related to consumer heterogeneity.<sup>3</sup>

Our paper focuses particularly on the effects of promotional activity on customers' reactions to future deals rather than future purchases. Previous work on this question has mostly used observational data with mixed results. Mela et al. (1997) use a discrete choice model with time-varying parameters and document that, in the long-run, price promotions are associated with heightened price sensitivity of both loyal and non-loyal customers. Jedidi et al. (1999) take advantage of the same long series analyzed by Mela et al. (1997) and show that promotions are associated with negative brand equity. Blattberg et al. (1995) suggest that increased promotions reduce the discount spike. In contrast (Zenor et al. 1998) indicate that they amplify it, while Bolton (1989) finds no effect. Boulding et al. (1995) suggest that the effect might vary by brand, and Narasimhan et al. (1996) indicate that the result might be category-specific. To the best of our knowledge, this paper is the first to provide a causal, strictly experimental estimate of the effects of current promotions on future promotion sensitivity.

We also contribute to the research on supply-side promotions by investigating the implications of heightened promotion sensitivity on promotional activities and equilibrium profits of competing manufacturers.<sup>4</sup> Our findings suggest that small firms benefit first from offering price promotions as unit margins increase. Moreover, heightened promotion sensitivity can increase the likelihood of promotions, potentially making all firms worse off. Whereas a monopolist could potentially benefit from heightened promotion sensitivity by offering deep discounts for a short period of time followed by a series of shallow discounts, rival firms instead race to the bottom until they find themselves in a state of fierce competition. Hence, we uncover a prisoner's dilemma that rationalizes the occurrence of promotional activities, despite their damaging effects. In particular, this finding is an instance of a Bertrand super-trap, as proposed by Cabral and Villas-Boas (2005), in which an apparent advantage for a monopolist (such as scope economies) can lead to all-across lower profits to competing firms over time.

Finally, our paper contributes to the literature on consumer search. While search behavior is usually unobservable in retail settings, we show that normalized preference parameters of the demand model induced by search can still be estimated by

<sup>3</sup> Anderson and Simester (2004) suggest our focus as a possible avenue for future research: "We cannot say how customers would have responded to [...] a subsequent discount. Investigating these issues would require different studies in which the experimental manipulations were [...] repeated in a subsequent catalog." In addition to solving the selection issue, our context and focus display other differences. In terms of the context, we work in different institutional settings (e.g., durable goods sold by catalog vs. non-durable goods sold in supermarkets), implying that the findings from one environment need not necessarily extend to the other. For example, the presence of immediate competitors in a limited amount of shelf space, or the limited amount of consumer learning in retail settings (Tuchman et al. (2017)), may lead to different forces and different results. Regarding the focus, we aim to analyze the implication of dynamic consumer behavior for the supply-side, as rationalized by consumers searching for deals, rather than focus on the extent to which different mechanisms may contribute to the main effect.

<sup>4</sup> See Raju (1992), Lal and Matutes (1994), and Freimer and Horsky (2008), and Villas-Boas and Villas-Boas (2008) for work considering the supply-side dynamics of price promotions, among others.

integrating over consumer search paths. Recovering the fundamental demand parameters allows us to investigate a precise demand mechanism, as well as to consider the supply-side equilibrium implications for competing firms. Moreover, the demand model not only makes the demand mechanism precise, but also lends itself naturally to allow complex dynamic forces, such as the impact of price promotions on consumers' beliefs and the resulting future demand.

We provide two additional results that can be useful for the estimation of consumer search models. In particular, we build on the seminal application by Kim et al. (2010), who estimate the sequential search model proposed by Weitzman (1979).<sup>5</sup> A critical step in estimating these models is the calculation of alternative-specific 'reservation values,' which are used to determine consumers' search sequences across products. We propose two alternative methods to the current interpolation approach. First, we show that in situations where the researcher is willing to consider agents face Logistic preference shocks, the reservation values are directly obtainable through a convenient closed-form expression. The Logistic distribution is attractive since it features unbounded support, and is often used as an analytical approximation of the Normal distribution.

Second, in situations where the researcher prefers a specific uncertainty structure, we prove a contraction-mapping result which implies that, in general, the reservation values are obtainable through fixed-point iteration. The method applies to a broad class of continuous distributions.

The remainder of the paper is organized as follows. The next section describes the experimental design, the data, and validates the experimental intervention. Section 3 presents the results of the field experiment. Section 4 describes the demand model, identification, estimation, and results. Section 4.4, in particular, presents the supply-side model and the counterfactual analyses. Section 5 concludes.

## 2 Field experiment

### 2.1 Experimental design

The experiment took place during the August-October period in 2013, and was conducted in partnership with a large retail firm in Chile, which holds 30% market share nationwide. The retailer's stores are organized into two retail sub-chains, which make use of different branding and perform separate marketing activities. In this paper, we report the experimental results of the intervention on the larger chain, where the depth of promotions of 170 top-selling products was manipulated across 10 stores, approximately one third of the stores held by the chain at the time.<sup>6</sup>

---

<sup>5</sup> See also Seiler (2013), Honka (2014), Bronnenberg et al. (2016) and Ke et al. (2016) for recent theoretical and empirical advances in modeling and understanding consumer search.

<sup>6</sup> While we performed the same experimental manipulation in both retail chains, we only report the results for the first chain, which is the larger one. The intervention did not produce statistically significant results for stores of the smaller chain. However, all results are directionally consistent across chains and dependent variables. We believe the smaller sizes of stores in the second chain are responsible for the absence of statistically significant effects. The results of the intervention in the smaller chain are available from the authors.

The intervention affected products in 17 categories: Beer, Bread, Breakfast Cereal, Candy, Cheese, Cold Cuts, Cookies, Cooking Oil, Fruit Juices, Meats, Milk, Pasta, Snacks, Soft Drinks, Tea, Water, and Yogurt.<sup>7</sup> In order to maximize the visibility of the intervention, we randomly selected 10 products (or sku's) from the subset of 15 products exhibiting the largest market shares in each category.<sup>8</sup>

In order to analyze the intertemporal response to varying promotion depths, we randomly assigned stores to one of two alternative conditions: (i) in "treated" stores, participating products featured deep discounts (approximately 30%) during the intervention phase (five weeks), and shallow discounts (approximately 10%) during the measurement phase (five weeks); (ii) in "control" stores, participating products featured shallow discounts (10%) during both experimental phases.<sup>9</sup> We verified with the retailer's management that the frequency with which individual sku's are experimentally promoted accords well with the usual practices of the retailer, and that the discount levels of 10% and 30% are near the lowest and highest discount levels typically offered by the retailer. We offer discounts on both conditions to isolate pure promotion effects from regular price effects.

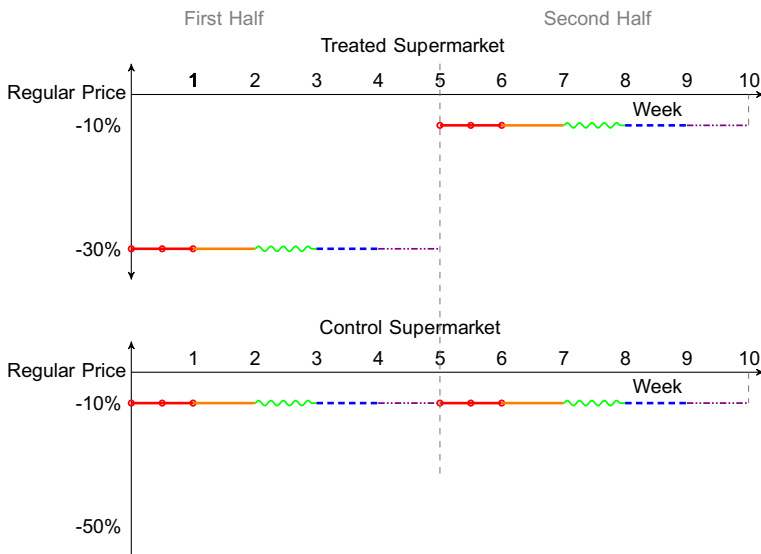
Following the retailer's standard practice, each product remained in the promotion condition for one week simultaneously across conditions. Products were placed on promotion on a Tuesday and remained in that state until Monday of the following week. We rotated the products to be promoted within each category weekly, in order to keep promotion frequencies consistent with the retailer's practice. This process is illustrated in Fig. 1. Each color/pattern represents a different pair of sku's within a given category. In each of the first five weeks of the experiment, the prices of two different sku's were marked down (approximately 30% in treated stores, 10% in control stores) with promotional tags. By the end of week five, the prices of all participating sku's had been intervened with. The promotion schedule was then repeated during the measurement phase, this time with equal markdowns of 10% in both types of stores. As we detail later, logistical reasons led us to opt for an approximate schedule to the one depicted in Fig. 1 during the second half of the experiment, in order to take the managers' logistical constraints into account.<sup>10</sup> The promotional schedule

<sup>7</sup>We detail the criteria used for category selection in Appendix A.

<sup>8</sup>Sku, or *stock keeping unit*, is a unique identifier of the product at the retailer.

<sup>9</sup>Regular prices at treated and control stores are very similar across store pairs since the retailer sets regular prices based on pricing "zones." The retailer defines pricing "zones" according to geographic locations, demographics and competition intensities. Each pair of treated/control stores always belongs to the same pricing zone.

<sup>10</sup>We were able to get buy-in for our manipulation during the first half of the experiment. However, the retailer viewed the initial promotion schedule as too heavy-handed to be implemented for the whole 10 weeks, and so an approximate alternative promotion schedule was agreed upon for the second half. The reasons for the changes were mainly due to previous agreements with manufacturers regarding sales targets and the planned promotion activity for our promoted products and their respective competitors. The changes were at a national level and were not related to local demand conditions. Moreover, they are constant across control/treated store pairs, as we further describe in Section 2.3, and so are unlikely to affect the validity of our estimates.



**Fig. 1** Timing of Discounts

was discussed explicitly with firm executives to ensure that the intervention was in line with current practice. As an illustration, in our 46 weeks of pre-experimental data, each product is promoted every 4.98 weeks on average, which is consistent with managerial accounts as well as the experimental frequency. We therefore believe the experimental findings do not result from consumers being exposed to unfamiliar stimuli.

Our experiment relies on pairwise randomization, i.e., the full sample of stores was divided into pairs according to covariate values, and each pair member was then assigned to the treatment or control group randomly (see Imbens (2011)). Out of all of the retailer's stores, we selected store pairs along three dimensions: (i) similar consumer demographics (age, gender, socioeconomic group); (ii) similar competition intensity; (iii) similar geographic location. The retailer provided us with store-level demographic information and with the classification of neighborhood factors to construct the measures. Geographic proximity ensures that seasonality does not play a role because treatment and control stores have the same holiday calendar and weather conditions. Finally, the tentative store assignments were discussed with the retailer to ensure that managerial knowledge was also taken into account. This resulted in the creation of six pairs of comparable stores which, as we explain later, yielded 5 usable store pairs. After the procedure was completed, each store in each pair was assigned to an experimental condition via coin flip.

Despite the large numbers of customers and products affected by the experiment, we acknowledge that the sample of 10 stores is limited in terms of the number of observations available for the analysis. Experimenting with prices across multiple stores, products and weeks has significant logistical and payoff implications

for firms. In fact, in our case, the 10 stores represent nearly a third of the total number of the chain's stores. Hence, we cannot necessarily rely on asymptotic results for the statistical analysis. Two significant steps were taken to address this issue. First, we introduce a finite-sample correction to the clustered standard errors in order to correct the potential downward bias induced by the small number of clusters. Second, we perform the main analysis on a sub-sample of matched consumers that are most comparable across treatment and control conditions. We detail these steps in Section 3.

## 2.2 Data

The primary dataset covers all transactions carried out by loyalty program members in the participating stores during the experimental period. The data include quantities purchased and the actual prices paid by a given customer for each sku across product categories. We track the purchases of a given consumer using identifiers from the retailer's loyalty program database. This loyalty program covers a substantial fraction of the retailer's total revenues (approximately 80%). We were also provided with cardholders' demographic information including their gender, age, and socio-economic group classification.

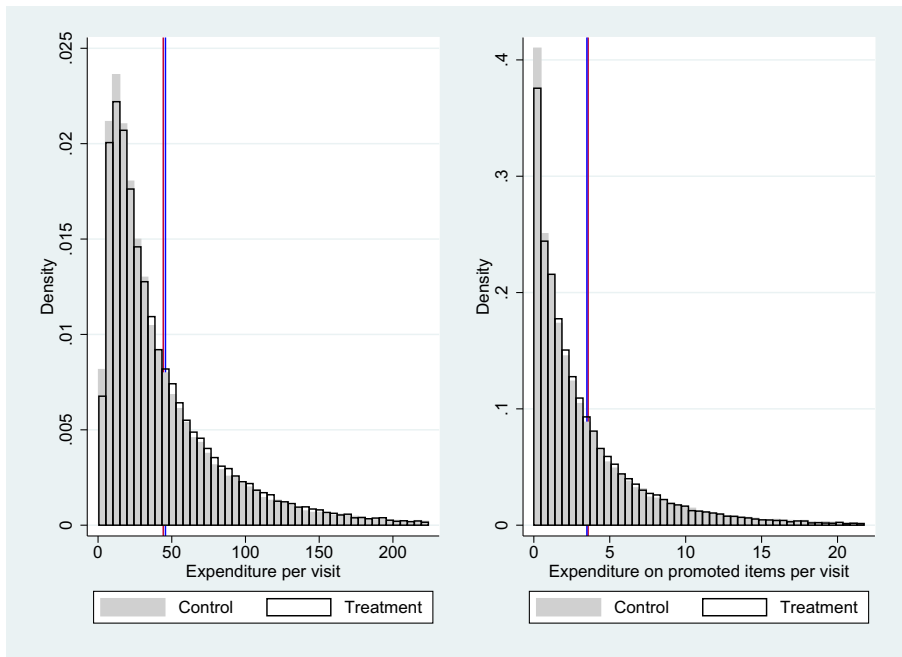
The analysis focuses on purchases of experimental goods by a sample of 234,063 loyalty program cardholders that meet the following conditions: 1) bought at least one (potentially non-experimental) sku in the top 31 retailer categories during the first half of the experiment and 2) did not visit more than one of the retailer's stores during the whole experimental period. The goal of the first condition is to focus on consumers who shop relatively frequently at the retailer, and who may have been exposed to the intervention. The second condition is imposed to guarantee that the same consumer was not exposed to different experimental conditions. Table 1 shows the descriptive statistics of the data during the measurement phase, per store-week

**Table 1** Descriptive Statistics by Week-Store

	Mean	Std. Dev.	Minimum	Maximum
Number of visits	10,267.6	3,742.7	3,239	17,323
Total sales (USD)	\$166,517.7	\$89,057.1	\$54,874.7	\$348,648.4
No. of items sold	70,636.7	33,588.7	20,218	130,831
– Experimental	1,707.3	944.7	375	4,548
– Non-experimental	68,929.5	32,735.2	19,783	127,259
Average basket size (items)	6.7	1.2	4.8	9.9
Number of weeks	5			
Number of stores	10			

Notes: The figures above are calculated during the second half of the experiment, for the selected sub-sample of 234,063 consumers, across experimental categories (excludes Candy, Cheese, Cooking Oil, and Meats categories)





**Fig. 2** Pre-experimental Shopping Behavior across Experimental Conditions

pair. On average, each store had about 10,000 customer visits each week and sold roughly 70,000 sku's weekly.<sup>11</sup>

The retailer made available a supplemental dataset containing individual-level behavior from the last week of September 2011 until the first week of August 2012, hence ending approximately one year before the beginning of the intervention period. The dataset contains individual-level behavioral statistics of 221,323 out of the 234,063 experimental consumers. Figure 2 compares the pre-experimental distributions of two measures of customer purchasing behavior: total expenditure per visit and spending on promoted categories per visit. We observe a high degree of overlap between the distributions of treated and control customers for both measures of pre-experimental shopping behavior. Furthermore, average pre-experimental expenditures in both promoted and non-promoted categories appear to be very similar across experimental conditions, as can be gauged from the proximity between the vertical lines, blue and red, depicting per-visit average expenditure for treated and control customers, respectively. We focus the analysis on the subset of 221,323 consumers to be able to control for historical information.

Table 2 presents further details on the comparison between treated and control customers. Individuals in both experimental conditions are highly similar along

<sup>11</sup>Throughout the paper, we use the symbol '\$' to denote US Dollars, which were calculated by converting from the local Chilean Peso currency, based on the exchange rate of 0.0016 USD/CLP obtained from Google on May 27th, 2015.

**Table 2** Descriptive Statistics by Experimental Condition

	Treated Stores (1)	Control Stores (2)
Demographics		
– Age	47.26 (14.26)	45.68 (14.27)
– Fraction Female	.657 (.475)	.648 (.478)
Pre-experimental expenditure (USD per visit)		
– Total Expenditure	\$66.70 (60.86)	\$55.43 (73.55)
– Expenditure on Experimental Categories	\$20.44 (21.43)	\$17.20 (22.16)
– Total Number of visits in 46 weeks	30.28 (12.92)	26.32 (12.97)
Number of customers	115,129	106,194

Notes: Standard deviations in parentheses. Pre-experimental data is available only for 221,323 out of the 234,063 individuals considered in the pooled regressions

demographic dimensions (age and gender), although relatively small differences exist at the behavioral level. These behaviors are controlled for in the main analysis and, as we explain later, we rely on these measures to subsequently focus the analysis on matched individuals.

### 2.3 Compliance with the experimental design

We actively engaged with the retailer to ensure the experiment was implemented according to plan. We first held a meeting with all the store managers to explain the importance of following the experimental protocol closely and maintaining an identical shopping environment in treated and control stores. Given that managers are often limited in the promotions they can offer, the experiment was generally well received even by managers of control stores. An executive from the retailer headquarters was named as the coordinator and supervisor of the experiment. This executive was in charge of ensuring that price lists were sent to the participating stores in a timely fashion. Finally, store managers were asked to write back to headquarters and submit photos of updated price promotions on a weekly basis, as depicted in Fig. 3.

Although 12 stores were initially selected to participate, only 10 stores fully took part in the experiment. The reason was that a manufacturer noticed the effects of the intervention, which apparently benefited a major rival in a specific product category. The concerns targeted two stores, which were immediately removed from the experiment. The two stores had originally been assigned to different store pairs and different conditions. However, the fact that these pairs shared the same classification based on customer demographics and competition level allowed us, with the agreement of management, to reassign their couples to a new store pair. This reassignment is unlikely to compromise the experimental design and consistency of point estimates.

To assess compliance, we compare the average price levels that were effectively implemented against the prices defined by the experimental design across treated and control stores. We perform the comparison for the implementation and measurement phases broken down by weeks and promotional status of the products.

Table 3 shows the comparison of the average prices during the intervention phase (first half) for the 17 categories involved in the experiment. The price comparison for products in the non-promoted and promoted weeks are in the leftmost and rightmost columns of the table, respectively. The average price difference of non-promoted products between treatment and control stores is  $-0.2\%$ , close to the ideal of the experimental design. Similarly, the average discount of promoted products was  $9.7\%$  and  $26.6\%$  in treated and control stores, respectively, relatively close to the ideal intervention of shallow and deep discounts.

While the results related to the intervention phase were generally deemed as satisfactory, concerns related to the Candy and Meats categories emerged. Table 3 shows that the Candy category was promoted at  $50\%$  discount levels in both treated and control stores and that the Meats category was promoted with discounts of nearly  $20\%$  in control stores. While we are not aware of the specific reasons for these deviations, we preferred to keep the analysis conservative by removing these potentially problematic categories. The reason is that, even if the phenomenon responsible for the deviations is exogenous, deviations from our intervention plan are counterproductive



**Fig. 3** Experimental Promotions in the Retail Space. Clockwise: Experimental Promotions in the Snacks, Tea and Cooking Oil Categories

**Table 3** Price Levels in the Intervention Phase

	No Promotion			Promotion			
	Control Store (1)	Treated Store (2)	% diff	Control Store (3)	Treated Store (4)	Disc. Control	Disc. Treated
Beer	\$6.63	\$6.60	−0.5%	\$6.25	\$5.11	5.7%	22.6%
Bread	\$2.30	\$2.30	0.1%	\$2.11	\$1.74	8.1%	24.4%
Breakfast cereal	\$3.25	\$3.24	−0.3%	\$2.84	\$2.37	12.5%	26.7%
Candy	\$3.58	\$3.58	−0.1%	\$1.83	\$1.52	<b>48.9%</b>	<b>57.4%</b>
Cheese	\$4.68	\$4.67	−0.3%	\$4.26	\$3.38	8.9%	27.5%
Cold cuts	\$6.50	\$6.43	−1.1%	\$6.47	\$5.13	0.5%	20.3%
Cookies	\$0.83	\$0.83	0.6%	\$0.75	\$0.62	8.9%	25.6%
Cooking oil	\$3.37	\$3.46	2.6%	\$3.21	\$2.64	4.8%	23.7%
Fruit juice	\$1.11	\$1.11	−0.2%	\$1.06	\$0.87	4.8%	22.1%
Meats	\$8.43	\$8.16	−3.2%	\$6.81	\$5.49	<b>19.2%</b>	32.7%
Milk	\$1.16	\$1.16	0.0%	\$1.11	\$0.88	4.4%	23.9%
Pasta	\$0.93	\$0.93	−0.3%	\$0.86	\$0.68	7.3%	26.3%
Snacks	\$2.03	\$2.03	0.4%	\$1.85	\$1.51	8.7%	25.7%
Soft drinks	\$2.02	\$2.02	−0.1%	\$1.98	\$1.60	1.9%	20.7%
Tea	\$2.67	\$2.66	−0.2%	\$2.54	\$2.09	4.7%	21.4%
Water	\$1.16	\$1.16	0.0%	\$1.08	\$0.87	6.8%	24.5%
Yogurt	\$0.34	\$0.34	−0.1%	\$0.31	\$0.26	9.2%	23.1%
Average			−0.2%			9.7%	26.6%

Notes: The table presents average prices at the category level during the first five weeks of the experimental period

in answering the research question of the differential dynamic effects of offering shallow vs. deep promotions.

Table 4 shows the comparison of the average prices during the measurement phase (second half) for the 17 categories involved in the experiment. As in the implementation phase, the difference between prices across control and treated stores for the non-promoted products is less than 1%. Equally satisfactory is the fact that the average discount of promoted products was 11.2% and 11.0% in treated and control stores, respectively. Although there exists some variation on the actual promotion level offered across categories, the differences in promotion levels are relatively small across types of stores. However, a few categories exhibit unexpected patterns: the Candy and Cheese categories feature extremely high discounts, and the Cooking Oil category was sold under promotion at above-average prices. This result led us to further remove the Cheese and Cooking Oil categories from the analysis.

Finally, we also compare prices across experimental halves. Table 5 shows the comparison of the average prices between the intervention and measurement phases

**Table 4** Price Levels in the Measurement Phase

	No promotion			Promotion			
	Control Store (5)	Treated Store (6)	% diff	Control Store (7)	Treated Store (8)	Disc. Control	Disc. Treated
Beer	\$6.92	\$6.95	0.4%	\$5.95	\$5.90	14.0%	15.1%
Bread	\$2.36	\$2.36	0.2%	\$2.22	\$2.19	5.9%	7.1%
Breakfast cereal	\$3.39	\$3.31	−2.4%	\$2.95	\$3.01	13.2%	8.9%
Candy	\$3.58	\$3.58	0.2%	\$1.93	\$1.90	<b>46.1%</b>	<b>47.1%</b>
Cheese	\$4.67	\$4.66	−0.2%	\$2.62	\$2.55	<b>43.8%</b>	<b>45.3%</b>
Cold cuts	\$6.82	\$6.69	−1.9%	\$6.08	\$6.05	10.8%	9.5%
Cookies	\$0.84	\$0.83	−0.8%	\$0.82	\$0.83	1.7%	0.5%
Cooking oil	\$3.60	\$3.62	0.3%	\$3.88	\$3.94	<b>−7.6%</b>	<b>−8.9%</b>
Fruit juice	\$1.18	\$1.17	−0.2%	\$1.02	\$1.02	12.8%	13.2%
Meats	\$8.56	\$8.22	−3.9%	\$8.24	\$7.97	3.7%	3.1%
Milk	\$1.21	\$1.22	0.2%	\$1.14	\$1.13	6.1%	7.0%
Pasta	\$0.94	\$0.94	−0.2%	\$0.88	\$0.87	6.2%	6.7%
Snacks	\$2.06	\$2.05	−0.3%	\$1.88	\$1.82	8.7%	11.2%
Soft drinks	\$2.05	\$2.04	−0.2%	\$1.91	\$1.87	6.5%	8.1%
Tea	\$2.61	\$2.60	−0.2%	\$2.54	\$2.50	2.6%	3.7%
Water	\$1.18	\$1.17	−0.5%	\$1.16	\$1.15	1.6%	1.8%
Yogurt	\$0.33	\$0.33	−0.5%	\$0.30	\$0.29	11.3%	11.2%
Average			−0.6%			11.0%	11.2%

Notes: The table presents average prices at the category level during the last five weeks of the experimental period

for the 17 categories. The price comparison for the non-promoted and promoted products are in the leftmost and rightmost columns of the table, respectively. The first two columns of Table 5 reveal that regular prices increased slightly across the experimental halves, with average increases of 2.3% and 1.8% in control and treated stores, respectively. Within-category differences fall within single digits, and as a result, we see no reason to eliminate observations based on these figures. The third column, however, reveals that promoted prices in control stores changed significantly in the Cheese, Cooking Oil and Meats categories. The expected change was equal to zero because products in control stores are supposed to be promoted at the 10% level throughout the experiment. Moreover, in the second half of the experiment, products in treated stores should also be promoted at 10%. The fourth column of Table 5 calculates the relative difference between promoted prices between treated stores in the second half, and control stores in the first half. As with the third column, the same three categories register higher than expected differences. The levels of regular and

**Table 5** Price and Promotion Differences across Experimental Phases

	% Change in regular prices		% Change vs. promotions in control store	
	Control [(5)-(1)]/(1)	Treated [(6)-(2)]/(2)	Control [(7)-(3)]/(3)	Treated [(8)-(3)]/(3)
Beer	4.4%	5.3%	-4.8%	-5.7%
Bread	2.5%	2.6%	4.9%	3.8%
Breakfast Cereal	4.5%	2.3%	3.7%	6.1%
Candy	-0.2%	0.1%	5.2%	3.5%
Cheese	-0.2%	-0.1%	<b>-38.5%</b>	<b>-40.2%</b>
Cold Cuts	5.0%	4.1%	-6.0%	-6.4%
Cookies	1.5%	0.1%	9.4%	9.9%
Cooking Oil	6.9%	4.5%	<b>20.8%</b>	<b>22.7%</b>
Fruit Juice	5.7%	5.7%	-3.3%	-3.9%
Meats	1.5%	0.8%	<b>20.9%</b>	<b>17.0%</b>
Milk	4.9%	5.2%	3.1%	2.2%
Pasta	0.8%	0.9%	2.1%	1.2%
Snacks	1.5%	0.8%	1.4%	-1.6%
Soft Drinks	1.3%	1.2%	-3.4%	-5.3%
Tea	-2.3%	-2.3%	-0.1%	-1.5%
Water	1.9%	1.4%	7.6%	6.9%
Yogurt	-1.0%	-1.4%	-3.2%	-3.6%
Average	2.3%	1.8%	1.2%	0.3%

Notes: The table presents average price differences (in percentage terms) across experimental phases using the respective columns in Tables 3 and 4

promotional prices in the remaining categories were accepted as complying with the experimental protocol, in light of the intervention complexity.<sup>12</sup>

### 3 Results

We analyze mean differences in consumer behavior between the control and treated stores during the measurement phase (second half) of the experiment. The main model aggregates store-category-SKU transactions to the individual level, in line with our interest in measuring the dynamic effects of promotions at the customer level. The main regression equation is given by

$$y_i = \alpha + \beta_1 T_i + \beta_2 X_i + \varepsilon_i \quad (1)$$

<sup>12</sup> All of the results of interest survive the inclusion of the problematic categories, although the significance of the treatment effects decreases. Moreover, a few confounding results arise, related to the unexpected significance of a few variables with no apparent explainable pattern.

where  $y_i$  is a customer behavior measure of interest in the measurement phase, and  $T_i$  is a dummy variable equal to one for consumers shopping in treated stores. Parameter  $\beta_1$  thus captures the change in  $y_i$  induced by exposure to deep discounts, rather than shallow, during the first half of the experiment. We include controls  $X_i$  for precision and robustness purposes, namely, store-pair fixed effects, age and age squared, gender, as well as two categorical variables of socioeconomic level provided by the retailer. We also introduce behavior-based controls available on the historical dataset provided by the retailer, namely, visit frequency, average weekly expenditure and average weekly expenditure on items in the experimental categories.<sup>13</sup> The analysis focuses on 7 behaviors of interest during the measurement half of the experiment: whether consumers are more likely to buy a promoted item (variable 1), the proportion (quantity and expenditure) of promoted vs. non-promoted items bought (variables 2 and 3), the quantities and expenditures on promoted and non-promoted items (variables 4-7). The first 3 variables are of main interest to sign the effects of the intervention, whereas the last 4 provide insights about the underlying behaviors that generate the results.

Two (related) notes are in order. First, given the nature of our setting, administering treatment conditions at the individual level is infeasible. The ideal experiment would randomly assign shoppers to a 30% or 10% discount first, and then track customers during the common second-stage. In this case, we would identify the average treatment effect statistically with simple regression analysis on the whole sample using robust standard errors. In a retail setting, however, random assignment at the customer level is infeasible, contrary to online settings, for example. Consider the possibility of providing in-store coupons to customers, in a randomized fashion. This procedure could likely introduce fairness concerns among individuals who did not receive the better coupons and it would not measure the effect of promotions in the way they are normally provided to consumers in stores. In our setting, randomization must take place at the store level, which naturally constrains the number of randomization units, especially given that the total number of stores is small at the population level, let alone the comparable stores used to form a sample. As a result, we do not assume that each consumer observation in regression (1) provides an independent draw from the population. Instead, shoppers at a given store may share observable and unobservable characteristics, making the data heteroscedastic at the observation level. Hence, each consumer's effective contribution to reducing standard error estimates is likely to be lower than in the i.i.d. case. We cluster standard errors at the store level to avoid downward biases in our standard error estimates (see Bertrand et al. (2004) for an in-depth discussion).

Second, the validity of standard error clustering relies on the asymptotic behavior of the estimator at the cluster, rather than at the individual, level. Given the relatively small number of stores/clusters available for the experiment, we introduce a finite correction to the standard errors. We implement the "wild bootstrap" procedure, as

<sup>13</sup>Our baseline specification pools across all purchases made by a given consumer, hence the use of a single sub-index  $i$  in Eq. 1.

proposed by Cameron et al. (2008), to correct for downward bias potentially induced in small samples.<sup>14</sup>

### 3.1 Analysis on full sample

Table 6 summarizes the results for each of the seven measures of interest. The first three columns summarize the results of interest, whereas columns 4–7 shed light on the underlying mechanism. All point estimates in columns 1–3 are positive, suggesting consumers exhibit heightened promotion sensitivity after being exposed to deep promotions. Treated consumers were more likely to buy promoted items (column (1)) than their control counterparts. They also have a higher share of promoted items in their baskets during the measurement phase of the experiment (columns (2) and (3)). In terms of the underlying forces, consumers exhibit higher purchase rates of all items, independently of promotional condition (columns (4)–(7)).

The point estimates suggest a positive treatment effect. Taken face value, these estimates mean that being exposed to deep promotions during the first half of the experiment led consumers to be on average 3.6 percentage points more likely to buy a promoted item during the second half, despite facing the same discount levels as their control counterparts. Using statistic  $E[\hat{y}_i | X_i, T_i = 1] \div E[\hat{y}_i | X_i, T_i = 0]$ , this effect translates to a relative increase of 22.4%. Despite this, the quantity effect is relatively modest: columns (2) and (3) suggest that exposure to deep promotions led consumers to shift 1.3% of items in their baskets towards promoted goods, which translates into a 4.7% relative increase.

Figure 4 decomposes the treatment effect across categories. The dependent variable is the indicator function (column (1) in Table 6) of whether the consumer made a purchase of a promoted good in the focal category. The point estimates range from 0.1% in the Tea Category up to 2.1% in the Milk category, with a number of them attaining significance at the 95% confidence level under the wild bootstrap standard errors. None of the category results support the hypothesis that deeper promotions lead to consumer deal addiction.<sup>15</sup>

The point estimates in Table 6, columns (4)–(7), further suggest that the shift in basket compositions appears to source from an overall increase in the total number of purchases. However, further inquiry is needed to assess the statistical robustness of the results. Rows (a)–(c) denote different approaches to calculating p-values of the estimated treatment effects. All coefficients are statistically significant at the 1% level when OLS standard errors are used. However, clustering standard errors (row b) decreases the significance of all treatment effect estimates, making results (1) and (3)

<sup>14</sup>See Cameron and Miller (2015) and Imbens and Kolesar (2015) for relevant discussions and practical guidance. Importantly, the wild bootstrap procedure is valid in the presence of right-hand indicator variables (see Cameron and Miller (2015), pp. 345–346). Previous versions of this manuscript employed bootstrap-t standard errors, which suffer from this limitation.

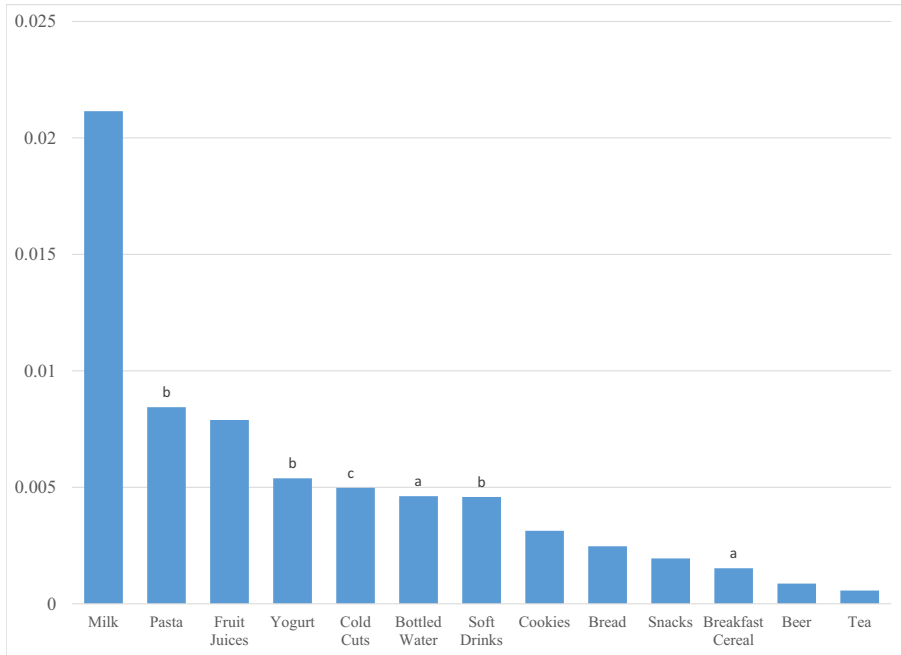
<sup>15</sup>Results on the heterogeneity of treatment effects are available from the authors. We find treatment effects are positively correlated with the number of promoted items bought in the past, controlling for average basket sizes. This finding is consistent with the idea that consumers who value promotions more, or equivalently, face lower costs when searching for deals, also show the most substantial increases in promotion sensitivities.



**Table 6** Effect of Treatment on Customer Behavior

	(1) Bought promoted sku	(2) % Items bought in promotion	(3) % Expen- diture on promoted sku's	(4) No. of promoted sku's bought	(5) Expenditure on promoted sku's	(6) No. of non- promoted sku's bought	(7) Expendi- ture on non- promoted sku's
Treatment							
(a): OLS std. errors	0.036 (0.000) <sup>c</sup>	0.013 (0.000) <sup>c</sup>	0.014 (0.000) <sup>c</sup>	0.233 (0.000) <sup>c</sup>	0.275 (0.000) <sup>c</sup>	0.658 (0.000) <sup>c</sup>	0.63 (0.000) <sup>c</sup>
(b): Clustered std. errors	(0.023) <sup>b</sup>	(0.063) <sup>a</sup>	(0.039) <sup>b</sup>	(0.09) <sup>a</sup>	(0.057) <sup>a</sup>	(0.06) <sup>a</sup>	(0.065) <sup>a</sup>
(c): Wild bootstrap	(0.132)	(0.212)	(0.164)	(0.314)	(0.258)	(0.268)	(0.26)
Constant	0.028 (0.416)	0.252 <sup>c</sup> (0.000)	0.243 <sup>c</sup> (0.000)	-0.48 (0.134)	-0.617 (0.138)	-1.443 (0.166)	-1.803 (0.126)
Controls & store group fixed effects	✓	✓	✓	✓	✓	✓	✓
R-Squared (Within)	0.082	0.003	0.003	0.071	0.071	0.124	0.141
N. Observations:	221,323	117,510	117,510	221,323	221,323	221,323	221,323

Notes: <sup>a</sup> $p \leq 0.10$ , <sup>b</sup> $p \leq 0.05$ , <sup>c</sup> $p \leq 0.01$ . p-values are in parentheses. Standard errors for p-values in (b) and (c) are clustered at the store level. p-values for the treatment effect in (c) and the constant are derived from the wild bootstrap procedure with 1,000 replications



**Fig. 4** Category-Level Treatment Effects. Notes: <sup>a</sup> $p \leq 0.10$ , <sup>b</sup> $p \leq 0.05$ , <sup>c</sup> $p \leq 0.01$ . p-values are in parentheses. Standard errors are clustered at the store level, and p-values for the treatment effect are derived from the wild bootstrap procedure with 1,000 draws. The analysis is based on matched customers.  $N=221,323$  in all regressions

significant at the 5% level, and the remaining ones significant at the 10% level. Row (c) shows the p-values calculated through the wild bootstrap procedure, which takes into account the potential bias introduced by the lack of asymptotic properties of the small number of clusters. In this case, none of the treatment effects are different than zero at the 90% confidence level.

It is worth discussing the potential sources for these results. One possibility is that our results are spurious. In other words, they would arise from the allocation of stores to experimental conditions, with the implication that another allocation outcome could have produced different qualitative results. Second, the context of our experiment is extremely noisy: supermarket environments are known to be crowded with marketing stimuli with the goal of affecting consumer behavior. These stimuli may limit the magnitude of the true treatment effects, as well as our ability to detect them reliably. This is a challenging problem to address, since increasing the number of experimental units (number of stores) is extremely costly, and administering treatment individually is not practical in this setting. In this case, our point estimates may be near the true ones, but the experiment may be underpowered. Third, it is also possible that the true significance percentiles fall somewhere between those in rows (a) and (c). While clustering is often used to eliminate spurious results obtained by overly-optimistic modeling assumptions, the procedure is also commonly described

as producing very conservative results and increasing the probability of type 2 errors. The next section investigates the power issues further.

### 3.2 Robustness

In order to investigate the statistical significance results further, we repeat the analysis on a subsample of consumers who are most similar on observable characteristics across experimental conditions. The underlying assumption is that the observable characteristics of these consumers are correlated with the unobserved ones. In this case, matching consumers on observable characteristics reduces the variance of the unobservable ones, resulting in a more efficient estimation of the treatment effects. This exercise is informative in that it predicts that estimation significance should increase for this case. Importantly, the matching procedure employed was not revisited after it was designed, and all analyses were performed after the completion of the matching procedure employed. This sequence of events ensures that the matching procedure is not contaminated by the results it generates, eliminating the potential for feedback effects and researcher bias. We outline the matching procedure and the characteristics of the matched sample in detail in Appendix B.2. We re-estimate Eq. 1 using the sub-sample of 26,964 customers who were matched across treatment and control conditions. As a reminder, we focus on customer behavior during the second half of the experiment in which promotion depths are held constant across experimental conditions.

In order to have a benchmark to compare our experimental results with, we first report the static effects of our intervention in Table 7. We find that consumers who face deep promotions are approximately 9% more likely to purchase a promoted item (column (1)), which translates to almost one additional promoted item bought, on average (column (4)). In addition, we find no statistically significant effect on the purchase behavior of items when they are not promoted.

Table 8 presents the results using the same matched subsample during the second half, when promotion activities are similar across treated and control groups, and confirms the result of heightened promotion sensitivity. The first column presents the effect of exposure to deep promotions on the purchase incidence of promoted products. Treatment induces an increase in the probability of buying a promoted sku in 5.3 percentage points and is statistically significant at the 95% confidence level. In relative terms, consumers exposed to deep promotions are 21.2% more likely to buy promoted items than their control counterparts, despite facing similar promotional levels. The second and third columns also suggest a shift in consumers' basket compositions towards promoted products, by more than 3 percentage points. Importantly, the results of the treatment effects of interest (columns (1)–(3)) are all significant at the 95% level. The differences between treated and control consumers also increase, suggesting that the most comparable consumers across experimental conditions may exhibit higher treatment effects.

Comparing the first columns of Tables 7 and 8 reveals that treated consumers exhibited 61% ( $0.053 \div 0.087$ ) of the initial promotion effect during the second half, at which time promotional activities were comparable across groups. That is, the magnitude of dynamic effects of price promotions, *ceteris paribus*, is 61% of the

**Table 7** Effect of Intervention on Customer Behavior - Matched Sample, First Half

	(1) Bought promoted sku	(2) % Items bought in promotion	(3) % Expen- diture on promoted sku's	(4) No. of promoted sku's bought	(5) Expenditure on pro- moted sku's	(6) No. of non- promoted sku's bought	(7) Expenditure on non- promoted sku's
Treatment	0.087 <sup>c</sup> (0.000)	0.062 <sup>c</sup> (0.000)	0.053 <sup>c</sup> (0.000)	0.873 <sup>c</sup> (0.008)	0.682 <sup>c</sup> (0.000)	0.585 (0.328)	0.366 (0.698)
Constant	0.015 (0.856)	0.215 <sup>c</sup> (0.004)	0.225 <sup>c</sup> (0.000)	-1.155 (0.23)	-1.382 <sup>a</sup> (0.086)	-2.598 (0.302)	-4.715 (0.118)
Controls & Store Group Fixed Effects	✓	✓	✓	✓	✓	✓	✓
R-Squared (Within)	0.098	0.017	0.016	0.073	0.066	0.143	0.11
N: Observations:	26,964	12,854	12,854	26,964	26,964	26,964	26,964

Notes: <sup>a</sup> $p \leq 0.10$ , <sup>b</sup> $p \leq 0.05$ , <sup>c</sup> $p \leq 0.01$ . p-values are in parentheses. Standard errors are clustered at the store level, and p-values are derived from the wild bootstrap procedure with 1,000 replications. The analysis is based on matched customers

**Table 8** Effect of Treatment on Customer Behavior - Matched Sample, Second Half

	(1) Bought promoted sku	(2) % Items bought in promotion	(3) % Expen- diture on promoted sku's	(4) No. of promoted sku's bought	(5) Expenditure on pro- moted sku's	(6) No. of non- promoted sku's bought	(7) Expenditure on non- promoted sku's
Treatment	0.053 <sup>b</sup> (0.034)	0.035 <sup>b</sup> (0.028)	0.032 <sup>b</sup> (0.036)	0.355 (0.184)	0.344 (0.146)	0.713 (0.284)	0.659 (0.142)
Constant	0.018 (0.754)	0.292 <sup>c</sup> (0.000)	0.276 <sup>c</sup> (0.000)	-0.562 (0.474)	-0.805 (0.412)	-1.894 (0.294)	-2.87 (0.172)
Controls & Store Group Fixed Effects	✓	✓	✓	✓	✓	✓	✓
R-Squared (Within)	0.085	0.008	0.008	0.067	0.066	0.119	0.142
N. Observations:	26,964	10,022	10,022	26,964	26,964	26,964	26,964

Notes: <sup>a</sup> $p \leq 0.10$ , <sup>b</sup> $p \leq 0.05$ , <sup>c</sup> $p \leq 0.01$ . p-values are in parentheses. Standard errors are clustered at the store level and p-values are derived from the wild bootstrap procedure with 1,000 replications. The analysis is based on matched customers

promotional effects induced by offering shallow (10%) vs. deep (30%) discounts during the intervention phase.

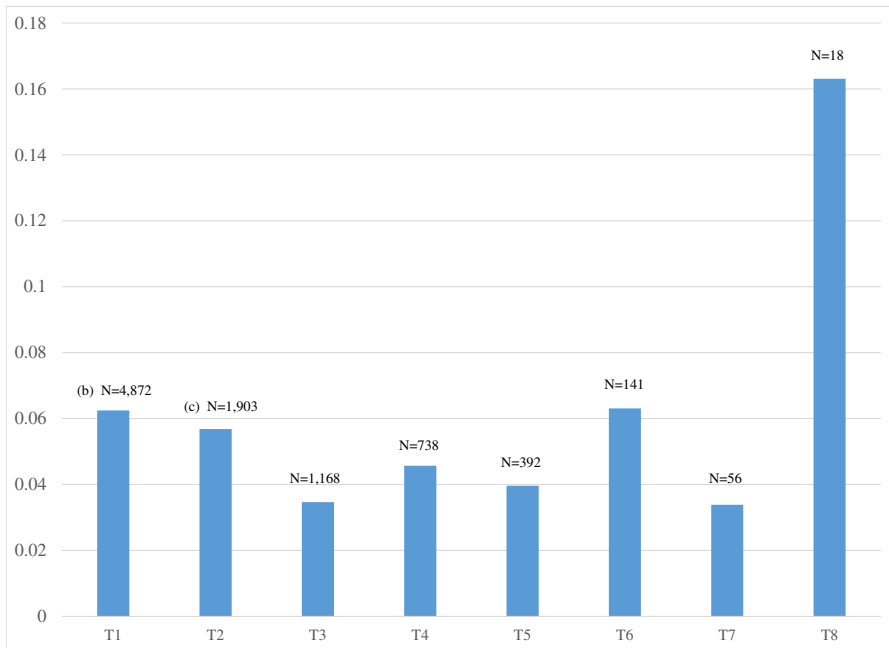
Comparing the point estimates in Table 6 to those in Table 8, we verify that there exist significant differences in the cases of columns 2 and 3.<sup>16</sup> The statistical significance in the matched sample is likely to derive from two sources: first, we have focused on a subset of consumers with similar observable characteristics. This procedure reduces the variance of the unobservables as long as these are correlated with the matched characteristics. A second factor that cannot be ruled out by the analysis is that these consumers also exhibit higher treatment effects, which are naturally easier to reject than small ones. Importantly, the results remain valid independently of the source of statistical significance.<sup>17</sup>

Figure 5 depicts treatment effects over time, by comparing the likelihood of treated and control consumers purchasing a promoted item during their first visit in the second half of the experiment. The lags in the “x” axis denote the number of weeks between consumers’ first visit during the second half of the experiment, and the last visit during the first half. We find positive treatment effects, with the first two lags being statistically different from zero and slightly higher than the estimated treatment effect in Table 8. The non-monotonic results are possibly due to the fact that the latter lags are very thin in terms of sample size (i.e., few consumers in our sample shop 6 and 7 weeks apart). A caveat to this analysis is that the probability of an individual falling into each bucket is not assured to be independent of her own treatment effect through experimental variation. Hence, the results should not be necessarily taken as causal, since they may be subject to endogeneity problems and/or may reflect heterogeneity in treatment effects along with the visit-frequency dimension.

We perform two main analyses to examine the robustness of our results. We first investigate to what extent the state dependence induced by our intervention could be affecting the results. State dependence, as operationalized by Guadagni and Little (1983) and further analyzed by Dube et al. (2010), can be generated by several psychological mechanisms, including learning (Erdem et al. 2003), consumers’ thinking costs (Shugan 1980) and consideration sets (Hauser and Wernerfelt 1990)). In our setting, treated consumers were exposed to deep promotions during the intervention phase, and so bought more experimental goods than their control counterparts at the time. Upon returning during the second half of the experiment, treated customers may

<sup>16</sup>We have simplified the comparison of coefficients by verifying whether the matched estimates fall within the confidence intervals of the unmatched ones. This method is used as an approximation, since the full analysis is complicated by the fact that the matched sample is necessarily correlated with the original one. Theoretically, without information about the joint distribution of the matched and unmatched coefficients, an implementation of non-nested tests à la Vuong (1989) is unavailable. While it is possible to use a bootstrap approach that performs the matching procedure on each bootstrap sample of the original dataset, the requirement of nesting the standard error correction implies a complex procedure as well as an impractical amount of computation time.

<sup>17</sup>There exists one limitation of this analysis, nonetheless, which is that the matching procedure was decided during the research program, and so even if the matching procedure was performed independently of empirical analyses, the matching sample represents an ‘additional draw’ made available to the researchers.



**Fig. 5** Treatment Effects Over Time. Notes: <sup>a</sup> $p \leq 0.10$ , <sup>b</sup> $p \leq 0.05$ , <sup>c</sup> $p \leq 0.01$ . In the x-axis, T# is the week difference between consumers' first visit during the measurement period and the last visit during the intervention period. Above each bar, the number of consumers in each bin. Standard errors are clustered at the store level, and p-values for the treatment effect are derived from the wild bootstrap procedure with 1,000 draws. The analysis is based on matched customers.  $N=18,172$

exhibit more purchases of promoted goods because of the positive state dependence generated by the purchase behavior in the first half.

The results indicate that state dependence is likely to play a minimal role. In particular, treatment effects are similar to the ones in the main analysis when we focus only on products bought for the first time during the measurement phase. Moreover, the statistical significance of the analysis on the matched sample also remains intact when the analysis focuses on products purchased initially during the second half of the experiment.

Second, we perform a placebo analysis by focusing on consumers who did not buy any products in the 31 main categories during the first half of the experiment, and so are less likely to have been exposed to the treatment. All point estimates decrease toward zero, often by one order of magnitude. Moreover, across different methods of calculating standard errors, only 2 out of the 21 p-values fall below the 10% threshold (compare with 14 out of 21 in Table 6; also, no p-values fall below 5% in the placebo analysis). Taken together, the results suggest that our main findings are robust to alternative explanations, in terms of both magnitude and precision.

In addition to these analyses, we also: 1) introduce additional promotion-related controls to the main regression; 2) use the implemented promotions, rather than treatment indicators, during the intervention phase as the focal treatment regressors;

3) discuss the possible role of stockpiling on our findings. In summary, all main results hold, and can sometimes be thought of as conservative. We present these analyses in Appendix B.

In the next section, we use the experimental findings to inform a model of consumer search, i.e., the experiment provides the parameter values at which we investigate the following two aspects. First, we ask whether consumer search can rationalize the experimental findings when historical promotional prices shed light on beliefs. Second, we investigate the implications of a demand curve induced by customers who search for deals to the profitability of competing manufacturers.

In the model, consumers may deliberate over product fit, which may change across shopping occasions due to fluctuations in needs, recollection of inventory at home, and other idiosyncratic circumstances. Product fit can be assessed through search, for example, as posited by Shugan (1980). Besides fit, search can also inform consumers about prices. As we now explain, in our setting, this applies primarily to promoted items. The reason is that, unlike promotional depths, changes to regular prices are fairly rare in our dataset. As a result, it does not appear sensible to have consumers search over the possible outcome of a constant value. Promotional depths, in contrast, do vary over products and time, and so can be learned by inspection. As with most supermarkets, in our setting products sold in promotion are identified with special-colored tags, which attempt to highlight the products on sale. We assume consumers observe these, and so can easily determine which are the items sold in promotion. However, additional inspection is required to assess the specific discount depths of the promoted items.

## 4 Searching for deals

Having documented the experimental finding that consumers exhibit heightened promotion sensitivity, we rationalize our evidence using a structural model of search behavior.<sup>18</sup> Since price expectations (or beliefs about promotion depth) are typically endogenous and search behaviors are usually unobservable in retail settings, we use our experimental setup and our individual-level data on purchases to inform a model in which consumers search for deals. The model captures a few essential features. First, it allows consumers to form beliefs about future promotion activities. In particular, the beliefs are informed by the historical promotion patterns at the retailer. Second, by allowing consumers to search for deals, it provides an explanation of consumer behavior that, together with the recovered beliefs, can rationalize our findings. Finally, it is used to assess the implications of a precise mechanism to firms competing through promotional activities.

---

<sup>18</sup>Consumer search has previously deserved attention in retail environments in the marketing literature (see for example Seiler (2013) and Seiler and Pinna (2017)).



## 4.1 Model

### 4.1.1 Consumer search

In order to derive implications about the phenomenon identified in the experiment, we consider a model in which consumers can search for deals before purchasing. We adopt the model of sequential search by Weitzman (1979), in which strategic consumers decide which products to evaluate, when to stop searching and which products to buy, if any. Evaluating alternatives is costly, and each consumer is required to incur search cost  $c > 0$  to consider each one.

To illustrate the search model, consider the case in which a consumer decides whether to evaluate a single product, with unknown utility  $u_1$ , while having in hand an alternative that provides known utility  $u_0$ . If the consumer searches (subscripts  $i$  and  $t$  are omitted for clarity), she expects to earn

$$U\text{Search}_1(u_0) = -c + Pr(u_1 \geq u_0) E[u_1 | u_1 \geq u_0] + Pr(u_1 < u_0) u_0 \quad (2)$$

which captures the fact that search may not necessarily be advantageous ex-post, but depends on whether  $u_1$  is higher or lower than the value of the current option,  $u_0$ . Relatedly, note that a consumer with a high value of  $u_0$  is less likely to search than one with a low value, ceteris paribus. Finally, there exists a level of  $u_0$  that makes the consumer indifferent between searching alternative 1 or not. Define this threshold as the reservation value of alternative 1, denoted by  $u_1^*$ , which is implicitly defined by the solution to equation  $U\text{Search}_1(u_1^*) = u_1^*$ .

Rewriting Eq. 2 for some generic product  $j$  yields its implicit reservation value, given by

$$z_j^* = -c + \int_{z_j^*}^{\infty} u dF_j(u) + F_j(z_j^*) z_j^* \quad (3)$$

where  $F_j(\cdot)$  is a product-specific cumulative distribution function.

Weitzman (1979) shows that optimal sequential search with multiple alternatives is characterized by a simple two-step rule: 1) calculate the reservation utility values  $u_j^*$  that make the consumer indifferent between searching each alternative  $j$  and not; 2) proceed to search alternatives by descending order of  $u_j^*$ , until the highest utility of the evaluated products is higher than all of the reservation values of the remaining uninspected alternatives, or alternatively, until all options have been inspected.

### 4.1.2 Consumer beliefs

Following the search literature, we assume consumers evaluate products in order to assess their idiosyncratic fit (e.g., products may be more useful at different occasions) as well as the prices they will have to pay to acquire them. Because in our setting promoted products are sold with special price labels made to stand out, we assume consumers can tell which products are sold in promotion easily. Moreover, regular prices rarely vary in our dataset, which informs the assumption that consumers know how much they will have to pay for products without special promotional labels a priori. Promotion depths do vary in our dataset, however, and as a result consumers

$$\begin{array}{c}
 t: \\
 \begin{array}{cc}
 & \begin{array}{cc} \text{Shallow} & \text{Deep} \end{array} \\
 \begin{array}{c} \text{Shallow} \\ \text{Deep} \end{array} & \begin{bmatrix} 1 - \omega^S & \omega^S \\ 1 - \omega^D & \omega^D \end{bmatrix}
 \end{array}
 \end{array}$$

**Fig. 6** Discount Depth Transition Probabilities for a Category

can also use search to identify the discount levels offered by promoted products sold with the promotional labels. Consistent with these assumptions, beliefs about current promotion depth only play a role in the evolution of prices conditional on a promotion occurring. Thus, we need only characterize transitions between low and high discounts, conditional on promotions taking place.<sup>19</sup>

We use historical data over promotion depths, as well as consumer visit behavior, to inform beliefs. Noticing that promotion depths vary over time, but are relatively highly correlated within each category during each promotional week, leads us to consider a category-level Markov process over category promotion depths, as depicted in Fig. 6.<sup>20</sup>

Notation  $\omega^\kappa$ ,  $\kappa \in \{S, D\}$ , indicates the probability of a category being promoted with a deep discount in visit  $t$ , given that it was promoted with a  $\kappa$ -level discount (Shallow or Deep) in visit  $t - 1$ . Under this specification, two consumers facing the same promotional activity may hold different beliefs over present discount depths, because of exposure to different promotional activities during their previous visits.<sup>21</sup> Empirically, we labeled discounts as shallow or deep based on the 15% cutoff, which falls between our experimental manipulations and also generates significant variation in the historical dataset.

An important note is that the matrix above governs promotion-depth transitions, *conditional* on items being sold in promotion. Given the assumption that consumers understand which products are sold at a regular price once at the aisle, adding an additional row and column with transitions from/to regular prices would be redundant in our setting. Rather, consumers keep track of the last promotion observed, as a predictor of the next one.

<sup>19</sup>In other words, consumers can distinguish between products being sold at regular prices and in promotion on shelves immediately because of the seller's different price tags. Given the lack of regular price variation, search informs consumers of idiosyncratic fit across all products, and also of promotional price levels for promoted products.

<sup>20</sup>It is possible that consumers can track promotion depths at the product level, especially for products they buy regularly. While promotion depths are fairly highly correlated within products of the same category at each point in time, we simplify the estimation by considering category-level beliefs. Note that we allow for beliefs at the individual product level in the counterfactual analysis. Finally, in the historical dataset, each category almost always has at least one product being promoted during each promotional period. We do not use the infrequent week-category combinations that do not feature any promotions in the estimation of the belief process.

<sup>21</sup>Index  $t$  is used to denote a consumer visit, except if noted otherwise.

### 4.1.3 Consumer utility

We assume consumer  $i$  derives utility from buying product  $j$  at purchase occasion  $t$  according to

$$u_{ijt} = \alpha_j + \beta SDep_{ijt} + \gamma^D d_{jt}^{Deep} + \gamma^S d_{jt}^{Shallow} + \varepsilon_{ijt} \quad (4)$$

where product fixed effect  $\alpha_j$  captures the time-invariant *net* utility of buying the product and paying its regular price. Regressor  $SDep_{ijt}$  captures whether individual  $i$  purchased product  $j$  in the previous visit (state dependence), and  $d_{jt}^{Deep}$  and  $d_{jt}^{Shallow}$  are indicators for the discount depths of product  $j$  on occasion  $t$ .<sup>22</sup> Finally,  $\varepsilon_{ijt}$  is an individual preference shock that captures the idiosyncratic fit of product  $j$  on occasion  $t$ . As explained before, we assume consumers know regular selling prices. Hence, when a consumer faces a product sold at a regular price, the taste shock is the only source of uncertainty. Consequently, for products sold at regular prices expression (4) can be re-written as

$$u_{ijt}^r = v_{ijt} + \varepsilon_{ijt} \quad (5)$$

where  $v_{ijt} \equiv \alpha_j + \beta SDep_{ijt}$  is the deterministic utility that consumer  $i$  knows she can derive from purchasing product  $j$  on occasion  $t$ .<sup>23</sup>

Since promoted products are featured with salient labels that invite consumers to search in order to assess whether the deals are attractive enough, consumers can immediately identify the promotional status of products but are required to inspect pricing labels to assess promotional values. Hence, promoted products have an additional uncertainty layer induced by the discount depths. Consumer  $i$ 's utility for a promoted product  $j$  is given by:

$$u_{ijt}^p = v_{ijt} + \varepsilon_{ijt} + \begin{cases} \gamma^D & \text{with probability } \omega_{ijt}^k \\ \gamma^S & \text{with probability } 1 - \omega_{ijt}^k \end{cases} \quad (6)$$

<sup>22</sup>Because in our experiment regular prices were not manipulated, and moreover because we observe little variation in past regular prices, we assume that  $p_{jt} = p_j$ . In the model,  $\alpha_j = \alpha'_j - \beta_0 p_j$ . Also, we simplify notation  $d_{jt}^{Deep}$  and  $d_{jt}^{Shallow}$  to indicate the type of discount product  $j$  was sold with to consumer  $i$  during her  $t^{th}$  purchasing occasion.

<sup>23</sup>Equations 4 and 5 present standard expected-utility maximizers. In contrast with reference-dependent models, where the utility specification depends on assumptions made on how reference prices are formed (see the meta-analysis in Briesch et al. (1997)), our model builds on minimal assumptions about consumer beliefs about prices. Moreover, given our experimental results and the nature of beliefs informed by historical promotion activities, reference-dependent models would predict results opposite to our experimental findings, due to treated consumers incurring losses when faced with shallow discounts during the second half of the experiment.

where  $\kappa \in \{S, D\}$  denotes the promotional status of product  $j$ 's category on occasion  $t - 1$ .

We assume  $\varepsilon_{ijt} \sim \mathcal{L}(0, 1)$ , where  $\mathcal{L}(\cdot)$  is the Logistic distribution, and explain the benefit of this choice in the next section. Consequently, the product utilities of regular priced and promoted products, conditional on the consumer  $i$ 's net utility from consumption and purchase history, follow distributions

$$u_{ijt}^r \Big|_{v_{ijt}} \sim F_{ijt}^r = \mathcal{L}(v_{ijt}, 1) \quad (7)$$

$$u_{ijt}^p \Big|_{v_{ijt}} \sim F_{ijt}^p = \omega_{ijt}^\kappa \mathcal{L}(v_{ijt} + \gamma^D, 1) + (1 - \omega_{ijt}^\kappa) \mathcal{L}(v_{ijt} + \gamma^S, 1) \quad (8)$$

Expression (7) states that, before search, the utilities from non-promoted products follow a logistic distribution. Expression (8) states that the utility of promoted products follows a mixture of Logistic distributions, induced by the unknown promotion depth consumers face before searching. Together with Eq. 3, distributions  $F_{ijt}^r$  and  $F_{ijt}^p$  determine the order in which alternatives are searched by consumer  $i$  on occasion  $t$ .

In our model, consumers differ because of idiosyncratic shocks as well as due to their states. Introducing additional heterogeneity would be helpful to match additional moments in the data. However, this would also complicate the objective functions of the supply-side significantly (see Zou (2014) for an interesting technique to tackle this issue in a monopoly setting). We refrain from considering counterfactuals that rely on consumer heterogeneity (e.g., targeting), and focus on analyses that depend on aggregate effects.

#### 4.1.4 Calculation of reservation values

Estimation involves solving non-linear Eq. 3 for each iteration/consumer/alternative/occasion combination. Since the task is time-consuming when tackled through numerical procedures, it is typically addressed by pre-computation of lookup tables of  $z_j^*$  as a function of  $c$  and  $v_{ijt}$ , under the assumption of normally distributed preference shocks (see Kim et al. (2010) and Honka and Chintagunta (2017)). The procedure involves a startup cost but is efficient for estimation of parameters in typical search contexts. Below we motivate the choice of Logistic preference shocks, which yields closed-form expressions for the reservation values:

**Proposition** (*logistic uncertainty*) *The reservation value in Eq. 3 admits a closed-form solution when  $F_j(u)$  is Logistic with location parameter  $v_{ijt}$  and scale equal to 1, namely  $z_j^* = v_{ijt} - \ln(\exp(c) - 1)$ .*

The result above is useful to recover the reservation values of regular-priced products in our context. However, it is not helpful to solve for the reservation values of promoted products, which induce a mixed distribution. The calculation of the reservation values for the promoted products is performed by use of the following result:

**Theorem** (contraction mapping) Function  $\Gamma(z) = -c + \int_z^\infty u dF_j(u) + F_j(z)z$  is a contraction mapping for any differentiable cumulative distribution function  $F_j(z)$  with finite moments  $E(u|u > z) \forall z \in \mathbb{R}$ .

The contraction mapping procedure can be used to recover the reservation values for a broad class of distributions while controlling for numerical precision in a parsimonious way.<sup>24</sup>

We prove both of the results above in Appendix C. These results provide additional options for the estimation of search-based models. When researchers face no restrictions on the particular uncertainty specification, the logistic distribution assumption yields a convenient closed-form solution to the reservation values. In settings where preference uncertainty is required to follow a different specification, however, the contraction mapping can be used to recover the reservation values efficiently for a general class of distributions. In our application, we face both cases, and so the calculation of reservation values for regular-priced products takes advantage of the first result, and the calculation for promoted products takes advantage of the second one.

## 4.2 Estimation

### 4.2.1 Search paths

At the core of our investigation is consumer choice, which can be characterized according to

$$Pr(Choice_{ijt} | \theta, X_{ijt}, \omega_{ijt}) \quad (9)$$

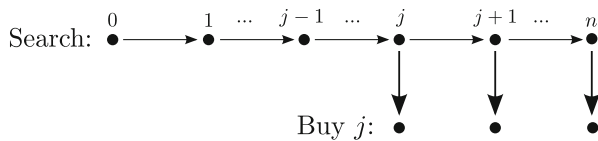
where  $Choice_{ijt}$  is the event that consumer  $i$  purchases product  $j$  on occasion  $t$ , conditional on some vector of preference parameters  $\theta = \{\alpha_1, \dots, \alpha_J, \beta, \gamma^D, \gamma^S\}$ , on some attributes  $X_{ijt} = \{SDep_{ijt}, d_{jt}^{Deep}, d_{jt}^{Shallow}\}$ , and on beliefs  $\omega_{ijt} = \{\omega_{ijt}^S, \omega_{ijt}^D\}$ . The object above is often investigated by use of a discrete choice model which, under typical assumptions, induces a logit likelihood that is especially amenable to empirical analysis. Since search paths are usually unobservable in retail settings such as ours, we integrate over the set  $\mathcal{S}$  of all potential search paths that may have led to the consumer's choice, according to

$$Pr(Choice_{ijt} | \theta, X_{it}, \omega_{it}) = \sum_{S_{it} \in \mathcal{S}} Pr(Choice_{ijt} | S_{it}, \theta, X_{it}, \omega_{it}) Pr(S_{it} | \theta, X_{it}, \omega_{it}) \quad (10)$$

The expression above takes into account that, depending on the consumer's preferences  $\theta$  over observables  $X_{it}$ , and on consumer beliefs  $\omega_{it}$ , different search paths  $S_{it}$  may arise with different probabilities.

For illustration purposes, consider the case of a consumer deciding between two goods with uncertain utilities  $u_1$  and  $u_2$ , and an outside option with known value  $u_0$ . Further, let product 1 have a higher reservation value than the second,  $z_1 > z_2$ ,

<sup>24</sup>Note that all three methods are relatively efficient in calculating reservation values. Hence, the choice of method depends on factors such as the specific search context and methodological convenience.



**Fig. 7** Search Paths Consistent with Choice of  $j$

given some consumer preferences and beliefs. In this case, a consumer may opt for the outside option as a result of different search paths: she may decide not to search, to search only the first alternative, or to search both. She opts for the outside option in each of the following search sequences:

$$\begin{aligned}
 \text{Path 1 : No search} & \quad \underbrace{u_0 > z_1}_{\text{stop search}} \\
 \text{Path 2 : Search 1 only} & \quad \underbrace{u_0 < z_1}_{\text{search 1}} \wedge \underbrace{u_0 > u_1}_{\text{prefer 0}} \wedge \underbrace{u_0 > z_2}_{\text{stop search}} \\
 \text{Path 3 : Search 1 \& 2} & \quad \underbrace{u_0 < z_1 \wedge u_0 > u_1}_{\text{search 1 \& prefer 0}} \wedge \underbrace{u_0 < z_2 \wedge u_0 > u_2}_{\text{search 2 \& prefer 0}}
 \end{aligned}$$

The consumer selects the outside option with probability  $Pr(\text{Path 1} \vee \text{Path 2} \vee \text{Path 3})$ , and in general, calculating the probability of each option being selected involves adding over the search paths consistent with the respective choices. Figure 7 below shows the potential search paths consistent with a choice of some alternative  $j$ .

In Fig. 7, searching an additional option corresponds to a lateral movement, and a downward one depicts the purchase of alternative  $j$ . In particular, for a consumer to be willing to search option  $j$  with reservation value  $z_j$ , she must have inspected options with higher reservation values before and have found that it was nonetheless worthwhile to continue to search, at least up to option  $j$ .

#### 4.2.2 Identification and estimation

We start by discussing the identification of the preference parameters. We focus the discussion on the product-specific constants,  $\alpha_j$ , since the identification of the preference parameters  $\beta$ ,  $\gamma^D$ , and  $\gamma^S$  then follows, given variation in observables.

Conditional on a search cost and consumer beliefs, changes to an alternative-specific constant  $\alpha_j$  only affect the reservation value of the same option  $j$ . Together with the assumption of mean zero valuation of the outside option, this fact implies a one-to-one correspondence with market shares.

As explained before, we estimate consumer beliefs using historical promotion depth transitions in the first stage. As for the search cost, in the absence of search data, it cannot be separately identified from the preference parameters. To see this, consider a market with two products, with market shares of 90% and 10%. Consumers may have an overwhelming preference for the first product because 1) it provides a much higher utility than the second one or 2) it is only slightly better than product two, but search costs are extremely high, such that most consumers prefer not

to inspect the second product. Given the normalization imposed on the search cost, preference parameters should be interpreted as relative to consumers' propensity to inspect alternatives.

We faced a few estimation challenges. First, the integration across search paths implies taking a large number of simulations. Second, the likelihood function featured a number of saddle points, which made identifying the global maximum challenging. Third, imposing a normalization on the search cost can affect the ability of gradient-based methods to find the preference parameters that rationalize relatively similar market shares in the data. We solve these issues by employing a global patterned search across the parameter values. In addition, we employed the smooth estimator proposed by McFadden (1989) to calculate the Hessian of the log-likelihood and the resulting standard errors. Our estimation procedure is performed in two stages. In the first stage, we estimate the consumer beliefs that are consistent with the promotion patterns observed in the pre-experimental dataset. In the second stage, we estimate the structural parameters conditional on the beliefs recovered in the first stage. The estimation details are provided in Appendix C.

We use both the intervention and the measurement phases to estimate the model at the consumer-visit level. The analysis focuses on the Milk category, as the implementation of the experiment closely followed the experimental design, as shown in Tables 3 and 5. Moreover, we focus on the behavior of customers of the matched subsample (see Appendix B.2) over both halves of the experiment, in order to improve estimation speed and efficiency. Within the Milk category, we model the actions of the largest three market share brands in the non-fat milk sub-category and assign purchases of all other types of regular (i.e. non-flavored) milk to the outside option. A number of factors support this sub-market choice. First, the three leading products belong to competing firms that hold a sizable total market share (33%) in the milk category and manage discounts independently with the retailer. Second, by including purchases of additional milk products (i.e. non-fat alternatives, 1.5 percent and whole milk offerings), our model rationalizes movements in intensive and extensive margins. However, we do not include occasions in which consumers bought no milk products, because promotions are likely to have a lower effect on overall consumption levels in this category.

Among the selected brands, brand 1 is the market leader and charges a price premium of 5% to 10% over the others, as shown in the first column of Table 9. Brand 2 is the retail chain's private label. Brand 3 holds the lowest market share and it is sold at an intermediate price point, as shown in the second column of Table 9.

**Table 9** Market Shares - Top 3 non-fat Milk Brands

	Avg. Price (Data, USD)	Actual Mkt. Share (Data)	Predicted Mkt. Share (Model)
Brand 1	\$1.10	15.08%	15.34%
Brand 2	\$0.99	11.49%	11.41%
Brand 3	\$1.04	6.77%	5.41%

Notes: Actual market shares computed from the data. Predicted market shares obtained through the structural model

**Table 10** Transition Matrix of Discount Depths

	(1) Shallow→Shallow	(2) Shallow→Deep	(3) Deep→Shallow	(4) Deep→Deep
Beer	0.636	0.364	0.444	0.556
Bread	0.973	0.027	0.667	0.333
Breakfast Cereal	0.769	0.231	0.500	0.500
Cold Cuts	0.882	0.118	0.571	0.429
Cookies	0.455	0.545	0.241	0.759
Fruit Juices	0.680	0.320	0.600	0.400
Milk	0.839	0.161	0.556	0.444
Pasta	0.947	0.053	0.500	0.500
Snacks	0.563	0.438	0.292	0.708
Soft Drinks	0.947	0.053	1.000	0.000
Tea	0.879	0.121	0.571	0.429
Water	0.886	0.114	0.600	0.400
Yogurt	0.975	0.025	N/A	N/A

### 4.3 Results

We recover the discount depth distribution from pre-experimental data to inform consumer beliefs of intertemporal promotion activities. We estimate the promotion depth Markov process at the category level using standard count methods. To compute transition probabilities, we assume an interpurchase time of 2 weeks, which accords well with the average time between visits that we observe in the data. The estimated transitions are later introduced into the search/purchase model at the individual level, depending on each consumer's visit patterns.<sup>25</sup>

We present the results of the discount-depth transition analysis in Table 10. With the exception of the Cookies category, consumers are most likely to find shallow discounts after being exposed to shallow discounts during their previous store visit. Moreover, deep discounts appear less “sticky” than shallow ones. Importantly, in all categories that exhibit a significant frequency of deep discounts, consumers are generally more likely to find deep discounts today after being exposed to deep discounts during their last visit (i.e. column 4 dominates column 2). This is consistent with the potential search explanation that consumers may be more willing to evaluate discounted alternatives after being exposed to deep discounts in the past. In contrast, the search mechanism proposed here cannot predict heightened promotion sensitivity if deals were negatively intertemporally correlated. In this case, treated consumers

<sup>25</sup>Our approach to recover consumer beliefs is consistent with most of the scanner panel literature. In contrast, imposing the Bayesian equilibrium concept would entail having consumers form beliefs based on firm-side fundamentals. Such equilibrium outcomes are much more challenging to investigate, as they require the researcher to 1) impose a rationale on the distributions of random variables affecting firms' willingness to offer promotions, and 2) later check possible promotional deviations based on each set of potential consumer beliefs.



**Table 11** Search Model Estimates

Parameter	Estimate
Alternative-specific Constants: $\alpha_1$	-1.344 <sup>c</sup> (0.056)
$\alpha_2$	-1.529 <sup>c</sup> (0.058)
$\alpha_3$	-1.875 <sup>c</sup> (0.058)
State dependence: $\beta$	3.227 <sup>c</sup> (0.103)
Shallow discount: $\gamma^S$	0.484 <sup>c</sup> (0.070)
Deep discount: $\gamma^D$	0.846 <sup>c</sup> (0.111)
N. Customers	26,964
N. Alternatives	3 per choice occasion+outside option

Notes: <sup>c</sup>p-value  $\leq 0.01$ .  
Standard errors in parentheses

should be less willing to search for deals, during the measurement phase, than their control counterparts. In other words, the recovered beliefs do not preclude the search mechanism as one of the explanations for our experimental findings. The recovered transition probabilities for the Milk category were incorporated into the model as consumer beliefs.<sup>26</sup>

Table 11 presents the estimates of the search model. As expected, all brand intercepts are negative, which reflects the fact that the outside option has the highest market share. Also, the state dependence parameter implies that purchasing a product today increases the purchase probability of the same product in the next occasion. The point estimates of both discount levels are positive, with the high discount coefficient being larger than the shallow discount coefficient, as expected. Table 9 shows that the model recovers the market shares well, with discrepancies being less than two percent between predicted and actual market shares.

## 4.4 Counterfactual analysis

### 4.4.1 Heightened promotion sensitivity

In order to test whether our search model can rationalize the experimental results, we compare the purchase probabilities of promoted goods for consumers who faced deep

<sup>26</sup>Because our intervention is relatively short, it is unlikely that it affected consumer beliefs at a fundamental level. Rather, the experimental effect is captured by the transition matrix. The recovered historical promotion transitions are used as the beliefs held by consumers for estimating the model, as well as in the counterfactual analysis.

discounts (treated consumers) with those who faced shallow discounts (control consumers). The model can rationalize our experimental evidence if treated consumers are more likely to buy promoted products than their control counterparts, despite facing the same discounts. We compute the difference in purchase probabilities between consumers who expect deep discounts relative to consumers who expect shallow discounts, despite facing the same shallow discount realization:

$$Pr \left( Buy_j \mid \underbrace{\omega^D}_{\text{Deep Belief}}, \underbrace{d_j^{Shallow} = 1}_{\text{Shallow Discount}} \right) - Pr \left( Buy_j \mid \underbrace{\omega^S}_{\text{Shallow Belief}}, \underbrace{d_j^{Shallow} = 1}_{\text{Shallow Discount}} \right) \quad (11)$$

We incorporate consumer beliefs in the following way. The first term of Eq. 11 corresponds to the probability of buying brand  $j$ , given that the consumer faced deep discounts in the past, and hence *expects* a shallow discount with 55.6% probability (according to Table 10, column 3). Similarly, the second term of Eq. 11 corresponds to the probability of buying brand  $j$  given that the consumer faced a shallow discount for milk in the past and hence *expects* a shallow promotion with 83.9% probability (Table 10, column 1).

The results are summarized in Table 12. They imply that the differential beliefs about promotion depths, as informed by the historical dataset, lead to a relative sales increase of 14.4% for promoted products, which can be compared with the result of 21.2% found in the analysis that takes state dependence into account (Appendix B). The underlying intuition is that, on the margin, consumers prefer to search products they believe are more likely to feature deep promotions. Hence, they display heightened promotion sensitivity despite effectively facing the same discount levels.

#### 4.4.2 Implications for competitors

We investigate the implications of consumer search for deals, as proposed by our model, for competing manufacturers. A naive interpretation of our results suggests that heightened promotion sensitivity should increase firm profits. For example, a firm may offer a deep discount once to induce search, and then only need to offer shallow discounts for a short period of time to generate positive results on sales.

**Table 12** Counterfactual Analysis: Heightened Promotion Sensitivity

	$Pr \left( Buy_j \mid \omega_{ic}^S, d_j^{Shallow} = 1 \right)$	$Pr \left( Buy_j \mid \omega_{ic}^D, d_j^{Shallow} = 1 \right)$	Relative Increase
Brand 1	11.48%	13.10%	14.15%
Brand 2	10.33%	11.85%	14.70%
Brand 3	7.96%	9.12%	14.57%
Average:	9.93%	11.36%	14.40%

Notes: The table presents predicted market shares for the three top brands in the non-fat milk sub-category for the case in which firms offer shallow discounts. Predicted market shares in column 1 assume the consumer expects a shallow discount with probability  $\Pr(\text{Shallow} \mid \text{Shallow})$ . Predicted market shares in column 2 assume the consumer expects a shallow discount with probability  $\Pr(\text{Shallow} \mid \text{Deep})$ . Column 3 is equal to  $((2)-(1))/(1)$

This interpretation is consistent with the previous counterfactual analysis, in which firms are believed to offer discounts unilaterally. We now consider the interaction of manufacturers who compete for consumers through promotion activities.

We study a two-period model in which firms compete through promotions strategically. Their action space is to sell at the regular price (i.e., no promotion), offer a shallow discount or a deep discount. Formally, in period  $t \in \{1, 2\}$ , firm  $j \in 1...n$  chooses a discount level  $Disc_{jt} \in \{None, Shallow, Deep\}$  and faces the demand function  $D_{jt}(Disc_{jt}, Disc_{-jt}, \Omega_t)$ , where  $Disc_{-jt}$  is a vector of discounts offered by firm  $j$ 's rivals in the same period, and  $\Omega_t$  is a vector of state variables relevant for competition. In particular, vector  $\Omega_t$  contains a sub-vector  $\hat{\omega}_t$  ( $n \times 1$ ) that summarizes consumers' expectations about the discount levels, and an additional vector with consumer state dependence information, related to the previous period's purchases. The first period's states,  $\Omega_1$ , are initialized by the researchers according to the brands' market shares in the data.

Each firm's objective is to maximize its expected two-period profit.<sup>27</sup> Equilibrium profits are given by

$$\Pi_j^* = \pi_{j1}^* + \pi_{j2}^* \quad (12)$$

where, in particular,  $\pi_{jt}^*$  is defined as

$$\pi_{jt}^* = \max_{Disc_{jt} \in \{None, Shallow, Deep\}} E \left\{ (p_j (1 - Disc_{jt}) - mc_j) \cdot D_j(Disc_{jt}, Disc_{-jt}^*, \Omega_t) \right\} \quad (13)$$

where  $p_j$  and  $mc_j$  are the firm's price and marginal cost, respectively. The demand side relies on the representative consumer's parameters estimated in the previous section. The regular prices in the data inform  $p_j$ , whereas we simulate the model at several degrees of marginal costs to incorporate the fact that price promotions may sometimes be directly linked to manufacturer-related conditions, including changes in opportunity costs of not selling.<sup>28</sup>

We focus on subgame perfect equilibria, such that firms understand that their actions influence future market conditions as well as the future actions of their competitors. The model is solved by backward induction. In addition, we use the fact that the firms' actions in the first period are a sufficient indicator to describe state vector  $\Omega_2$ .<sup>29</sup>

Once the second-period payoffs are calculated for each of the first-period action profiles, a deviation analysis is performed in order to isolate potential equilibria of

<sup>27</sup>We do not discount second-period payoffs, given the weekly timing of promotions.

<sup>28</sup>In our context, most promotions are initiated by manufacturers. We assume 100% passthrough in the analysis, although the effective rate is slightly lower. While we condition on different opportunity costs, in reality, they are likely to change over time, leading to different equilibrium promotion profiles in different weeks.

<sup>29</sup>Theoretically, firms could observe individual consumers' buying behaviors, and hence have a precise idea of the realized state dependence in the market at the individual level. In reality, firms may not always take all consumer behavior into account, and thus have to integrate over the distribution of state dependence in order to optimize their promotional offerings. In our case, our homogeneous demand structure allows us to summarize state dependence effects through past market shares.

the second subgame. Tentative equilibrium path profits are then plugged into the first-period optimization problem, where all action profiles are re-visited. We identify all pure subgame perfect equilibria by inspecting all possible first-stage action profiles. When a pure-strategy equilibrium does not exist, an exhaustive search is employed to uncover mixed strategy outcomes.<sup>30</sup>

We compare the market competition with and without heightened promotion sensitivity. In the first scenario, consumers search according to the recovered demand-side estimates and recovered beliefs for the Milk category. In the second scenario, consumers expect shallow discounts with constant probability, never adapting their beliefs about promotion activities. Specifically, in the first scenario consumers share the same initial belief of facing shallow discounts with 83.9% chance, but decrease it to 55.6% in the second period for firms who offer deep discounts in the first period. In contrast, in the second scenario, consumers expect shallow discounts occur with probability 83.9% in both periods, independently of firms' promotional activities. Hence, the baseline beliefs are constant across the scenarios, but firms can affect consumer beliefs in the first one. Firms can sell their products at regular prices, equal to the ones in the dataset, at shallow 4.4% discounts or deep 23.9% discounts, in accordance to the discounts effectively implemented in the experiment, as reported in Table 3.

We characterize market outcomes at a wide range of marginal costs in order to contemplate the following cases: 1) in some circumstances retailers want to promote specific products, making the manufacturer's relevant counterfactual margin higher than just the strict contribution margin; 2) manufacturers face inventory and cost variations over time as well as expiration dates, all of which affect their opportunity cost of selling. Such forces can lead accounting margins to underestimate economic ones. In the counterfactual analysis, a higher margin can be interpreted as a higher incentive to offer promotions.

Table 13 presents the subgame perfect equilibrium action profiles at different margin levels, when consumers exhibit heightened promotion sensitivity, or alternatively, use promotion transitions to form their beliefs. In the first period, firms prefer to engage in shallow promotions, up to the 40% margin level. From the 50% level onward, firm 3 transitions to offering deep promotions. These actions are in contrast with those in the second period, in which all firms prefer shallow promotions. The reason for this difference is that, in the first period, offering deep promotions can be beneficial because firms can then take advantage of state dependence in the second period. The policy change by firm 3 is intuitive in the sense that it is the firm with the lowest market share, and as a result, the one with the most to gain in inviting consumers to include it in their search paths.

When margins are high (80%), or alternatively, when opportunity costs of not selling are very high, the effective cost of offering deep promotions is lower, and so

<sup>30</sup>Given the number of players and the size of the action space, finding mixed strategy equilibria can be tedious. Whenever a pure strategy equilibrium was not found, our estimation code prepared a file to be read by the Gambit software ([www.gambit-project.org](http://www.gambit-project.org)), which was then used to run all algorithms available to generate an exhaustive list of mixed strategy outcomes. These algorithms always found a unique mixed-strategy outcome.

**Table 13** Competition Counterfactuals: Heightened Promotion Sensitivity

Firm \ Margin	30%	40%	50%	60%	70%	80%	90%
a) Subgame Perfect Equilibrium:							
Period 1							
Firm 1	<i>S</i>	<i>S</i>	<i>S</i>	<i>S</i>	<i>S</i>	{ <i>S</i> : 71%; <i>D</i> : 29%} <i>D</i>	
Firm 2	<i>S</i>	<i>S</i>	<i>S</i>	<i>S</i>	<i>S</i>	{ <i>S</i> : 49%; <i>D</i> : 51%} <i>D</i>	
Firm 3	<i>S</i>	<i>S</i>	<i>D</i>	<i>D</i>	<i>D</i>	<i>D</i>	<i>D</i>
Period 2							
Firm 1	<i>S</i>	<i>S</i>	<i>S</i>	<i>S</i>	<i>S</i>	<i>S</i>	<i>S</i>
Firm 2	<i>S</i>	<i>S</i>	<i>S</i>	<i>S</i>	<i>S</i>	<i>S</i>	<i>S</i>
Firm 3	<i>S</i>	<i>S</i>	<i>S</i>	<i>S</i>	<i>S</i>	<i>S</i>	<i>S</i>
b) Equilibrium Expected payoffs, per <i>potential</i> customer (assumes at most 1 purchase per customer), 2 weeks (USD): $\pi_{j1}^* + \pi_{j2}^*$							
Firm 1	\$ 0.11	\$ 0.16	\$ 0.19	\$ 0.24	\$ 0.28	\$ 0.32	\$ 0.36
Firm 2	\$ 0.08	\$ 0.11	\$ 0.14	\$ 0.17	\$ 0.20	\$ 0.23	\$ 0.26
Firm 3	\$ 0.05	\$ 0.06	\$ 0.08	\$ 0.11	\$ 0.13	\$ 0.14	\$ 0.15

all firms prefer to offer deep promotions, albeit sometimes only probabilistically: at the 80% level, firms 1 and 2 engage in mixed strategies between shallow and deep promotions.

Table 14 presents the same figures for the case in which consumers do not exhibit heightened promotion sensitivity. As before, all firms offer shallow discounts in the second period. In the first period, however, they prefer to offer discounts in fewer conditions. Firm 3 only offers deep discounts starting at the 60% margin level, and the remaining firms only begin offering deep discounts, probabilistically, at the 90% margin level. The reason for the increased reluctance to introduce deep discounts, when compared to the scenario of heightened promotion sensitivity, is that consumers do not expect a higher likelihood of a firm promoting in the second period as a function of its promotional activity in the first period. As a result, current promotions no longer provide an incentive for consumers to search in the second period, and so the returns of offering deep discounts in the first period become lower.

Table 15 shows the profit ratios between both scenarios, where a number above 1 means that the firm makes higher profits in the heightened promotion sensitivity scenario. In the 50%-70% margin range, firm 3 is better off offering a deep discount. The rationale is that, on average, firm 3 is the last one to be searched, and so it is the one that gains the most from inducing search by consumers through discounts. Firm 3's discount has a -4% profitability impact on firm 1, at the 50% margin, which then attenuates at higher margin levels. At 80% margins, firms 1 and 2 offer deep discounts with positive probability. The result is that all firms become worse off because of heightened promotion sensitivity. Cabral and Villas-Boas (2005) have coined this outcome as a Bertrand supertrap, in which an apparent advantage for a monopolist ends up effectively decreasing all firms' profits due to fiercer competition.

**Table 14** Competition Counterfactuals: No Heightened Promotion Sensitivity

Firm \ Margin	30%	40%	50%	60%	70%	80%	90%
a) Subgame Perfect Equilibrium:							
Period 1							
Firm 1	<i>S</i>	<i>S</i>	<i>S</i>	<i>S</i>	<i>S</i>	<i>S</i>	{ <i>S</i> : 68%; <i>D</i> : 32%}
Firm 2	<i>S</i>	<i>S</i>	<i>S</i>	<i>S</i>	<i>S</i>	<i>S</i>	{ <i>S</i> : 64%; <i>D</i> : 36%}
Firm 3	<i>S</i>	<i>S</i>	<i>S</i>	<i>D</i>	<i>D</i>	<i>D</i>	<i>D</i>
Period 2							
Firm 1	<i>S</i>	<i>S</i>	<i>S</i>	<i>S</i>	<i>S</i>	<i>S</i>	<i>S</i>
Firm 2	<i>S</i>	<i>S</i>	<i>S</i>	<i>S</i>	<i>S</i>	<i>S</i>	<i>S</i>
Firm 3	<i>S</i>	<i>S</i>	<i>S</i>	<i>S</i>	<i>S</i>	<i>S</i>	<i>S</i>
b) Equilibrium Expected Payoffs, per <i>potential</i> customer (assumes at most 1 purchase per customer), 2 weeks (USD): $\pi_{j1}^* + \pi_{j2}^*$							
Firm 1	\$ 0.11	\$ 0.16	\$ 0.20	\$ 0.24	\$ 0.28	\$ 0.32	\$ 0.36
Firm 2	\$ 0.08	\$ 0.11	\$ 0.14	\$ 0.17	\$ 0.20	\$ 0.23	\$ 0.26
Firm 3	\$ 0.05	\$ 0.06	\$ 0.08	\$ 0.10	\$ 0.12	\$ 0.15	\$ 0.16

The results indicate that, in competition, promotional pressure sources from firms with smaller market shares first. Moreover, firm 3's gains from dynamic beliefs are non-monotonic in contribution margins: at moderate margins, it benefits from being able to offer deep discounts, generating heightened promotion sensitivity in the second period. However, once more prominent firms also offer deep discounts, firm 3 becomes worse off than if it faced myopic consumers. Heightened promotion sensitivity can thus benefit small firms when they are the only ones promoting by inducing additional search. However, once the whole category finds it beneficial to promote, all firms may become worse off, with small firms losing the most.

**Table 15** Firms' Profitability Implications of Heightened Promotion Sensitivity

Firm \ Margin	30%	40%	50%	60%	70%	80%	90%
Firm 1	100.0%	100.0%	96.0%	99.8%	99.8%	98.8%	98.7%
Firm 2	100.0%	100.0%	99.8%	99.7%	99.7%	97.7%	100.0%
Firm 3	100.0%	100.0%	101.9%	102.6%	102.6%	95.2%	92.5%

Notes: The table presents the ratio of the firms' profits under the assumption that consumers exhibit heightened promotion sensitivity relative to the absence of dynamic effects

## 5 Conclusion

In this paper, we have reported the results of a large scale field experiment that involved changing the prices of 170 products across 17 categories in 10 supermarkets, in order to investigate the dynamic effects of price promotions in a retail setting. We found that deep promotions heighten customers' future promotion sensitivities. In particular, customers were 22.4% more likely to buy promoted goods after being exposed to 30% discounts rather than 10%, *ceteris paribus*. Along the same lines, the proportion of promoted goods in consumers' baskets increased by 4.7% for treated consumers relative to their control counterparts. The point estimates of the treatment effects are positive across all categories.

We perform the analysis on a matched subsample to verify the robustness of our results. Consistent with the hypothesis that the main analysis is underpowered, the statistical significance of the treatment effect increases in the matched subsample. The procedure also provides statistical power to parse the experimental effects from those potentially generated by state dependence. Finally, a placebo test shows no significant effects on consumers who did not purchase from a major category during the first half of the experiment.

We use the experimental findings to investigate a model in which consumers search for deals, and show that, conditional on the historical promotion patterns observed at the retailer, the model can rationalize our data. The underlying intuition is that when consumers believe that promotion depths are positively correlated over time, offering a promotion today can be viewed as an invitation to consider the same product tomorrow. The counterfactual analysis documents that as the incentive to offer promotions increases, smaller brands with lower market shares have a higher incentive to offer promotions to invite consumers to consider them.

Our results provide a rationale for why some managers complain about competing through promotions, while at the same time we observe deep promotions being routinely offered. Since consumers exposed to deep promotions display heightened promotion sensitivity, and furthermore, since firms can induce greater subsequent purchases by offering deep promotions, competition on the intensity of promotional activities can hurt firms' profitability: firms may find themselves in a prisoner's dilemma in which they compete fiercely on promotional activities.

Regarding future research opportunities, it may be interesting to consider how heightened promotion sensitivity affects firms in the long run, as well as the effects on retailer profitability. Although by assuming full passthrough our model does not focus on the retailer's role, an implication is that the retailer can use passthrough as a coordination device that allows firms to move to an outcome characterized by fewer intensive promotional activities and greater overall profitability. Since larger margins provide an incentive for firms to offer deeper promotions, it is likely that the retailer can induce firms to reduce their reliance on deep promotions by lowering the extent to which trade promotions are passed through to retail prices, much like a central bank 'heats up' or 'cools down' an economy.

Another idea is to consider identifying own vs. cross-category effects. Parsing these effects would require varying the profile of promoted categories across stores, which entails a more demanding experiment. In fact, it is possible that our results

capture these two sources, some of the effects being partly due to the fact that we manipulated promotions across multiple categories simultaneously. Such a research agenda would be novel for the field and especially useful for practitioners.

## **Appendix A: Criteria for category selection**

We selected categories with the goal of providing the maximum amount of useful variation. First, we wanted to limit the influence of stockpiling behavior on the response to the promotion stimulus. If consumers respond to promotions by anticipating purchases, then severe post promotion dips could affect our estimates. On this basis, we excluded a few categories for which households' inventory costs were deemed to be very low (e.g., soups) and others for which consumers could keep the product in inventory for a period of time, well beyond the post-promotion period (e.g., coffee). A second related consideration for including a category was the length of the typical interpurchase time observed in the category. In particular, we excluded those categories for which typical interpurchase times exceeded 5 weeks on average. Third, we only included categories that had already been promoted on a regular basis. Since our focus is on the effects of changes in promotion depth, we wanted to keep the frequency with which products were placed on promotion as constant as possible. This requirement led us to exclude categories such as "baked goods" which were rarely, if ever, placed on promotion. Fourth, we included categories that were purchased across different demographic segments (i.e., heterogeneous in terms of socioeconomic groups and ages). By imposing this requirement, we wanted to ensure that the same categories would be relevant across all stores included in the experimental design. Fifth, we chose categories in which consumers were unlikely to use the presence of promotion as an input in their assessment of a product's quality. For instance, the presence of promotions in specific categories (e.g., fresh produce) can be interpreted as a negative quality signal, e.g., the product is about to expire or does not sell well, and the promotion is seen as an attempt to sell it rapidly. Sixth, we chose categories with different degrees of brand loyalty, e.g., soft drinks are well-known for having a few star brands with very loyal consumers, whereas milk exhibits more generic products, likely to be considered close substitutes by more consumers. Other considerations that played a role in our choice of categories were avoiding categories in which stockouts were known to occur more frequently and avoiding categories with a small number of brands.

## **Appendix B: Robustness/alternative mechanisms**

### **B.1 Continuous treatment variable**

One of the consequences of the experimental protocol not having been implemented by store managers exactly as designed is that different stores effectively received different treatment intensities (see Table 3). Therefore, it is useful to verify whether



the same results hold when the empirical treatment, rather than the theoretical one, is introduced into the econometric specification. We consider two approaches. First, we calculate the average promotional product price for each store during the first half, across the experimental goods. Using this statistic as the treatment variable takes into account that consumers who shop at different stores may have been effectively exposed to different treatment intensities.

Table 16 summarizes the results (even-numbered columns) and compares them with the main regression results based on the treatment indicator variable (odd-numbered columns). We find that an *increase* in the discount depth (i.e., lower prices through promotions) during the intervention phase leads to a *higher* purchase likelihood during the second half of the experiment. Consumers are 10.4% more likely to buy a promoted item during the second half of the experiment for each dollar received through discounts during the first half. The same qualitative results are found in terms of the composition of consumers' basket, in terms of the percentage of items and of expenditure.

In order to provide further robustness of our results, we also perform the same analysis at the category level. In this case, the regressor of interest is the average promotional price across experimental goods of each category. In this case, an observation is an individual/category combination. Because most individuals do not buy from most categories, this leads to many zeros on the left-hand side.

The results are presented in Table 17. All results remain statistically significant, and consistent with the original directions, although a slight loss in power may have been induced by the number of zeros introduced by this approach. Overall, our results are robust to using the empirical treatment intensities.

## B.2 Improving efficiency through matching

We complement our analysis with a matching procedure. Much like the use of control regressors, a matching method reduces the variance of the unobserved error term by taking advantage of the correlation between observable and unobservable characteristics. This approach shifts the emphasis from the cluster to the *individual* level by matching, within pairwise randomized stores, those individuals who have balanced covariates before the experimental period (Rubin 1973, 1979; Imbens and Rubin 2015).

In our case, the matching technique allows us to construct a large subsample of *statistical twins* (one twin buying at a control store and the other at a treated one) to ensure identical pre-treatment purchasing behavior between treatment and control *at the individual level*.<sup>31</sup> Notice that in our setting, the exogeneity of the treatment is guaranteed by the experimental design and matching is only needed to identify similar customers based on historical data.<sup>32</sup>

<sup>31</sup>This matching procedure is standard when randomization is at the cluster level, but statistical analysis is at the individual level (Imbens 2011).

<sup>32</sup>Unlike our paper, the matching technique is typically used with non-experimental data in order to balance relevant pre-treatment covariates between treatment and control groups (Imbens and Rubin 2015).

**Table 16** Effect of Treatment on Customer Behavior - Intervention Prices

Dep. Variable:	(1)	(2)	(3)	(4)	(5)	(6)
	Bought promoted sku		% Items bought in promotion		% Expenditure on promoted sku's	
	Treatment dummy	Avg. Price during Intervention	Treatment dummy	Avg. Price during Intervention	Treatment dummy	Avg. Price during Intervention
Treatment	0.053 <sup>b</sup> (0.034)	-0.104 <sup>b</sup> (0.02)	0.035 <sup>b</sup> (0.028)	-0.069 <sup>c</sup> (0.004)	0.032 <sup>b</sup> (0.036)	-0.063 <sup>b</sup> (0.016)
Constant	0.018 (0.754)	0.29 <sup>a</sup> (0.084)	0.292 <sup>c</sup> (0.000)	0.473 <sup>c</sup> (0.000)	0.276 <sup>c</sup> (0.000)	0.44 <sup>c</sup> (0.000)
Controls & Store Group Fixed Effects	✓	✓	✓	✓	✓	✓
R-Squared (Within)	0.085	0.09	0.008	0.008	0.008	
N: Observations:	26,964	26,964	10,022	10,022	10,022	10,022
Avg. Promo Price across Stores (Control)		\$2.66		\$2.66		\$2.66
Avg. Promo Price across Stores (Control)		\$2.13		\$2.13		\$2.13

Notes: <sup>a</sup>  $p \leq 0.10$ , <sup>b</sup>  $p \leq 0.05$ , <sup>c</sup>  $p \leq 0.01$ . p-values are in parentheses. Standard errors are clustered at the store level, and p-values are derived from the wild bootstrap procedure with 1,000 replications. The analysis is based on matched customers. (Odd columns repeated from Table 8.) In columns (2), (4), and (6), the treatment regressor is the within-store average promotional price faced by consumers for the experimental items during the intervention period, depending on the stores where they shop

**Table 17** Effect of Treatment on Customer Behavior - Average Prices by Category

Dep. Variable:	(1)	(2)		(3)		(4)		(5)	(6)
	Bought promoted sku	Treatment dummy	Avg. Price	Category	% Items bought in promotion	Treatment dummy	Avg. Price	Category	Expenditure on promoted sku's
Treatment	0.053 <sup>b</sup> (0.034)		-0.002 <sup>b</sup> (0.044)		0.035 <sup>b</sup> (0.028)		-0.004 <sup>b</sup> (0.034)		0.032 <sup>b</sup> (0.036)
Constant	0.018 (0.754)		0.005 <sup>a</sup> (0.578)		0.292 <sup>c</sup> (0.000)		0.288 <sup>c</sup> (0.000)		0.276 <sup>c</sup> (0.000)
Controls & Store Group Fixed Effects	✓		✓		✓		✓		✓
R-Squared (Within)	0.085		0.01		0.008		0.002		0.008
N: Observations:	26,964		350,532		10,022		35,760		10,022

Notes: <sup>a</sup>  $p \leq 0.10$ , <sup>b</sup>  $p \leq 0.05$ , <sup>c</sup>  $p \leq 0.01$ . p-values are in parentheses. Standard errors are clustered at the store level, and p-values are derived from the wild bootstrap procedure with 1,000 replications. The analysis is based on matched customers. (Odd columns repeated from Table 8.) In columns (2), (4), and (6), the treatment regressor is the within-store/category average promotional price faced by consumers for the experimental items during the intervention period, depending on the stores where they shop

### B.2.1 Matching framework

To construct our sample of statistical twins, we introduce a recent matching technique developed by Zubizarreta (2012). This matching technique takes advantage of new developments in optimization to match individuals in multiple aspects, which until recently was an unfeasible task due to the large dimensionality of the problem. Matching individuals on several dimensions encompasses other matching techniques, such as propensity score, by creating a superior and easily interpretable matching sample.

Formally, let  $\mathcal{T} = \{t_1, \dots, t_T\}$  be the set of treated units, and  $\mathcal{C} = \{c_1, \dots, c_C\}$ , the set of potential controls. Without loss of generality, suppose  $T \leq C$ . Each treated unit  $t \in \mathcal{T}$  has a  $P$  dimensional vector of observed covariates  $\mathbf{x}_t = \{x_{t,1}, \dots, x_{t,P}\}$ , and each control  $c \in \mathcal{C}$  has a similar vector  $\mathbf{x}_c = \{x_{c,1}, \dots, x_{c,P}\}$ . Let the assignment indicator  $a_{t,c}$  be equal to 1 if treated unit  $t$  is assigned to control  $c$ , and 0 otherwise; and denote the entire assignment matrix by  $\mathbf{a}$ .<sup>33</sup> The optimal assignment problem is given by:

$$\min_{\mathbf{a}} \sum_{t \in \mathcal{T}} \sum_{c \in \mathcal{C}} \delta_{t,c} a_{t,c} \quad (14)$$

$$\text{subject to } \sum_{c \in \mathcal{C}} a_{t,c} = 1, \quad t \in \mathcal{T} \quad (15)$$

$$\sum_{t \in \mathcal{T}} a_{t,c} \leq 1, \quad c \in \mathcal{C} \quad (16)$$

$$a_{t,c} \in \{0, 1\}, \quad t \in \mathcal{T}, c \in \mathcal{C} \quad (17)$$

$$\left| \sum_{t \in \mathcal{T}} \sum_{c \in \mathcal{C}} \frac{x_{c,j} a_{t,c}}{T} - \bar{x}_{\mathcal{T},j} \right| \leq \varepsilon_j, \quad j \in \{1, \dots, P\} \quad (18)$$

where  $\delta_{t,c} \in [0, \infty)$  is a distance function between treated and control units (e.g. Euclidean distance),  $\sum_{t \in \mathcal{T}} \sum_{c \in \mathcal{C}} \frac{x_{c,j} a_{t,c}}{T}$  denotes the average covariate  $j$  of assigned controls and  $\bar{x}_{\mathcal{T},j}$  denotes the average covariate  $j$  across all treated individuals.

The goal of the matching program is to minimize the total sum of distances between treated units and matched controls as stated in expression (14). The first three constraints describe the integer nature of the assignment problem: Each treatment unit is paired with one control unit (Eq. 15) and not all control units should be used (Eq. 16).<sup>34</sup> The set of constraints given by expression (18) introduces an upper bound on the difference allowed between treatment and control individuals for each covariate, according to  $\varepsilon_j > 0$ , the pre-determined tolerance level for covariate  $j \in \{1, \dots, P\}$ . This last set of constraints is a distinctive feature of the mixed-integer programming (MIP) matching approach proposed by Zubizarreta (2012).

<sup>33</sup>The assignment problem is a mixed-integer programming (MIP) problem where some of the decision variables are constrained to be integer values at the optimal solution.

<sup>34</sup>Note that, for the assignment problem, the labeling of treatment and control units is irrelevant for its optimal solution. We relabel some stores to increase the sample size of the matched sample.

**Table 18** Universe of Potential Pairs

Sample:	Before matching		After matching		
	(1)	(2)	(3)	(4)	(5)
Store pair	Treatment	Control	Treatment	Control	Sample
1	32,711	8,535	1,647	1,647	3,294
2	20,477	30,746	5,647	5,647	11,294
3	16,035	39,773	4,116	4,116	8,232
4	20,230	17,497	1,322	1,322	2,644
5	16,741	18,578	750	750	1,500
Total	106,194	115,129	13,482	13,482	26,964

Notes: Individuals matched by the Mixed Integer Programming procedure (Zubizarreta 2012) using pre-experimental data. Pre-experimental data is available only for 221,323 out of the 234,063 individuals considered in the pool regressions

### B.2.2 Matching supermarket customers

We use the individual-level data available through the retailer's loyalty card to match pairs of control and treated consumers based on the pre-experimental records according to the procedure described above. To construct statistical twins, we consider demographic and behavior-based covariates: Age, gender, the weekly average of total expenditure, the weekly average of total spending in the experimental categories, and the frequency of trips to the store. We then applied the criteria that customers are required to buy an item in at least one of the 31 main categories during the first half of the experiment and not visit multiple stores during the experimental period. The historical dataset used for this task covers a 46 week period, ending approximately one year before the experiment.

Consumers' historical behavior toward promoted goods is a natural variable to include in this analysis.<sup>35</sup> In our case, the supermarket chain does not hold records that allow us to identify promotions during the pre-experimental period directly. We construct the promotional variable for the pre-experimental period using the algorithm proposed by Kehoe and Midrigan (2015). This algorithm identifies regular prices in a posted price time series as the modal price within a rolling window of five weeks centered on the current week.<sup>36</sup> We define a promotional event for a given price trajectory as a period when the posted price for a given UPC-store lies below the regular price (for the same UPC-store) derived from the Kehoe and Midrigan (2015) price filter.<sup>37</sup> We use the depth measurements to create two new variables: the share of promoted items bought by each consumer during the historical dataset, and the number of promoted items bought per consumer-visit during the same period.

<sup>35</sup>We thank an anonymous referee for this suggestion.

<sup>36</sup>See the Appendix in Kehoe and Midrigan (2015) for a detailed description of the method.

<sup>37</sup>Despite its merits, one limitation of this procedure is that, due to the data requirements involved, it does not provide a complete sequence of promotion depths for all products in all stores in all periods.

**Table 19** Pre-treatment Covariates of Control and Treated Matched Individuals

Pre-treatment covariate	Control	Treatment	Difference	p-value
Average weekly total expenditure (USD)	\$79.71	\$80.08	−\$0.38	(0.14)
Average weekly expenditure on experimental categories	\$26.09	\$26.31	−\$0.22	(0.30)
Age	47.26	47.33	−0.07	(0.07)
Fraction female	0.66	0.64	0.03*	(0.00)
Total number of visits in 46 weeks	27.27	28.11	−0.83*	(0.00)
Number of promoted items over total items	0.0023	0.0018	0.0005	(0.10)
Number of promoted items per visit	0.0044	0.0044	0.0000	(0.92)

Top 5 covariates sample size: 26,964. Bottom 2 covariates sample size: 26,872

Table 18 presents the sample sizes of the universe of customers before matching and those who were matched by the MIP matching procedure. Columns (1) and (2) present the number of customers who faced the experimental promotional activity in each store pair and for whom we have historical data. Overall, as shown in columns (3), (4) and (5), the MIP matching generated 13,482 one-to-one customer pairs (one control and one treated customer), distributed across 10 stores of our retail chain. Table 19 presents the resulting covariates of the final matched sample.<sup>38</sup> The last column reports the p-value for the null hypothesis of identical means. We obtain close averages in total expenditure, expenditure in promoted categories, and age between treatment and control matched individuals. Given the large sample size of individuals, the tests reject nearly same means in gender and number of trips, although the table shows that the actual values are quite similar. As for the purchase behavior metrics related to promoted goods (last two variables in in Table 19) added later per the review team's suggestion, we do not reject that those variables have the same means across conditions. Moreover, the absolute differences for these variables are slim, if anything, favoring the control group towards being more active in purchasing promoted goods (thus making our results potentially conservative).

In order to further test the effects of the pre-experimental promotion-related variables, we include the historical share of promoted products bought as well as the historical mean number of promoted products purchased per visit in the main analysis as control variables. We report the results in Table 20. We find virtually no differences to the treatment effects (compare to Table 8). We interpret these findings, also in connection with those reported in Table 19, as the matching procedure having performed well in matching consumers overall, through the five dimensions mentioned above, such that the remaining dimensions had little effect on the treatment estimates.

<sup>38</sup>We explored different values for the tolerance parameters  $\varepsilon_j$  for each covariate  $j \in \{1, \dots, P\}$ , to account for the trade-off between the proximity measures of the paired customers and the resulting sample size. On the one hand, large values of  $\varepsilon_j$  lead to poorly matched pairs, while on the other, smaller values of  $\varepsilon_j$  reduce the sample size. In fact, some combinations of small values of  $\varepsilon_j$  imply no feasible solutions, i.e., no assignment meets the desired levels of balance on covariates. Importantly, no outcome analyses were performed during this stage.

**Table 20** Effect of Treatment on Customer Behavior with Promotion-Related Controls - Matched Sample

	(1) Bought promoted sku	(2) % Items bought in promotion	(3) % Expen- diture on promoted sku's	(4) No. of promoted sku's bought	(5) Expenditure on pro- moted sku's	(6) No. of non- promoted sku's bought	(7) Expenditure on non-promoted sku's
Treatment	0.054 <sup>b</sup> (0.032)	0.035 <sup>b</sup> (0.028)	0.032 <sup>b</sup> (0.03)	0.358 (0.214)	0.347 (0.182)	0.715 (0.318)	0.664 (0.304)
Constant	0.005 (0.992)	0.288 <sup>c</sup> (0.000)	0.27 <sup>c</sup> (0.000)	-0.634 (0.406)	-0.88 (0.328)	-2.159 (0.244)	-3.182 (0.13)
Controls & Store Group Fixed Effects	✓	✓	✓	✓	✓	✓	✓
R-Squared (Within)	0.085	0.008	0.008	0.067	0.066	0.12	0.143
N. Observations:	26,872	10,012	10,012	26,872	26,872	26,872	26,872

Notes: <sup>a</sup>  $p \leq 0.10$ , <sup>b</sup>  $p \leq 0.05$ , <sup>c</sup>  $p \leq 0.01$ . p-values are in parentheses. Standard errors are clustered at the store level, and p-values are derived from the wild bootstrap procedure with 1,000 replications. The analysis is based on matched customers

### B.3 State dependence

To understand how state dependence could explain our results, consider a pair of similar customers who only differ on the experimental condition they were exposed to in the intervention phase. Assume that the treated customer bought a promoted product because of the deep discount, whereas the control customer decided not to purchase it, given its lower promotional discount (10%). It is possible that, during the second half of the experiment, both customers visited the store on the week that the same product was on promotion once again, at a shallow level. In this case, the treated customer may be more likely to buy the product in the second half, not because of heightened promotion sensitivity, but rather because of state dependence.<sup>39</sup>

We repeat the main analysis, but now only consider, for each consumer, purchases of goods that were not bought during the first half of the experiment. While this procedure is expected to mechanically decrease the treatment effect due to ignoring relevant data, it has the merit of parsing out the effect of state dependence. Table 21 summarizes the results for the full sample: the treatment effect in column (1) remains statistically insignificant, and all treatment effect estimates decrease slightly. However, statistic  $E[\hat{y}_i | X_i, T_i = 1] \div E[\hat{y}_i | X_i, T_i = 0]$  produces an estimate of a 21.2% relative increase of purchases of promoted products, which is similar to the one of 22.4% found before, when state dependence was not controlled for.<sup>40</sup>

Given the slight changes in significance and minimal changes to treatment effect estimates, we believe the new results are due to an overly-stringent test rather than state-dependence being responsible for the results. In order to investigate this issue further, we consider the same analysis on the matched sample, discussed in the previous section, and present the results in Table 22. All results of interest (columns (1)–(3)) remain statistically significant in this case, with the treatment estimates falling only slightly. Taken together, the results imply that state dependence may play a role in our measurement, but is unlikely to be responsible for the finding of heightened promotion sensitivity.

### B.4 Stockpiling behavior

Our results are likely to be made conservative by consumer stockpiling behaviors. The reason is that treated consumers are likely to hold higher inventories than control ones, due to the exposure to deep promotions during the intervention phase. As

<sup>39</sup>To be clear, the effect of state dependence can go either way: customers may also return in a week where their previously purchased product is no longer promoted, buy it (due to state dependence), and as a result, buy fewer (other) promoted goods during the second half of the experiment.

<sup>40</sup>It is worth clarifying the way the sample underlying these estimates was constructed. Following our choice of including in the analysis purchases (and non-purchases) by all consumers who visited a store and made a purchase in the first-half of the experiment, we included in the above estimation the purchase decisions of all individuals who made a purchase in the first-half of the experiment regardless of whether they bought a “new” product in promotion during the second-half. In the estimation reported above, purchases of the products previously purchased during the first half of the experiment were included as zeros in this analysis.



**Table 21** Effect of Treatment on Customer Behavior for New Purchases

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	Bought promoted sku	% Items bought in promotion	% Expenditure on promoted sku's	No. of promoted sku's bought	Expenditure on promoted sku's	No. of non-promoted sku's bought	Expenditure on non-promoted sku's
Treatment	0.031 (0.136)	0.014 (0.146)	0.015 (0.126)	0.137 (0.292)	0.178 (0.232)	0.324 (0.254)	0.306 (0.26)
Constant	0.034 (0.25)	0.258 <sup>c</sup> (0.000)	0.25 <sup>c</sup> (0.000)	-0.152 (0.354)	-0.255 (0.292)	-0.376 (0.364)	-0.419 (0.344)
Controls & Store Group Fixed Effects	✓	✓	✓	✓	✓	✓	✓
R-Squared (Within)	0.05	0.002	0.002	0.038	0.038	0.079	0.081
N. Observations:	221,323	107,235	107,235	221,323	221,323	221,323	221,323

Notes: <sup>a</sup> $p \leq 0.10$ , <sup>b</sup> $p \leq 0.05$ , <sup>c</sup> $p \leq 0.01$ . p-values are clustered at the store level, and p-values are derived from the wild bootstrap procedure with 1,000 replications

**Table 22** Effect of Treatment on Customer Behavior for New Purchases - Matched Sample

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	Bought promoted sku	% Items bought in promotion	% Expenditure on promoted sku's	No. of promoted sku's bought	Expenditure on promoted sku's	No. of non-promoted sku's bought	Expenditure on non-promoted sku's
Treatment	0.046 <sup>b</sup> (0.05)	0.034 <sup>c</sup> (0.000)	0.031 <sup>c</sup> (0.008)	0.226 (0.11)	0.22 (0.112)	0.364 (0.27)	0.324 (0.382)
Constant	0.056 (0.102)	0.312 <sup>c</sup> (0.004)	0.292 <sup>c</sup> (0.000)	-0.084 (0.764)	-0.225 (0.558)	-4.199 (0.312)	-0.705 (0.27)
Controls & Store Group Fixed Effects	✓	✓	✓	✓	✓	✓	✓
R-Squared (Within)	0.049	0.008	0.007	0.037	0.032	0.074	0.08
N. Observations:	26,964	8,330	8,330	26,964	26,964	26,964	26,964

Notes: <sup>a</sup> $p \leq 0.10$ , <sup>b</sup> $p \leq 0.05$ , <sup>c</sup> $p \leq 0.01$ . p-values are in parentheses. Standard errors are clustered at the store level, and p-values are derived from the wild bootstrap procedure with 1,000 replications. The analysis is based on matched customers

a result, stockpiling is expected to dampen treated consumer purchases during the second half of the experiment, including those of promoted products.

## B.5 Placebo test

In this section, we introduce a placebo test designed to assess whether our experimental intervention is likely to be effectively responsible for the differences in consumer behavior across treated and control pools. We focus the analysis on customers who did not visit the supermarket, or alternatively, did not use their loyalty cards during the first half of the experiment. Since these customers are less likely to have been exposed to the differential treatment conditions, we expect to find lower magnitudes and very few statistical significance of treatment effects (i.e., less than 5% of measured effects should be significant).

The results of the analysis are presented in Table 23 and are quite different from those in Table 6. First, only two behaviors are found to be marginally significant across different dependent variables and methods of calculating standard errors. Second, all point estimates fall below the original ones, often by an order of magnitude. If taken ‘as is’, the point estimate in column (1) would set a lower bound on the real treatment effect of 0.024, or a 10.7% increase  $\left(0.024 \div E \left[ \hat{y}_i^{full\ sample} \mid X_i, T_i = 0 \right] \right)$  in relative terms.

These results are positive in the sense that consumers who are less likely to have been exposed to the intervention exhibit lower treatment effects, generally without statistical significance. The results are reassuring in terms of assessing experimental validity, but not without limitations. In particular, the choice of not purchasing from a major category during the first half of the experiment is unlikely to be exogenous. Because of this, the lack of a statistically significant treatment effect can source from selection, i.e., this test may sample from consumers who respond less to price promotions in the first place, and who exhibit lower promotion sensitivity effects as well. However, a countervailing force is the fact that it is impossible to rule out that some of these customers were exposed to our intervention but purchased without using their loyalty card during the first half of the experiment. This force would contaminate the results in the opposite direction, making the treatment effects in Table 23 conservative (in other words, the analysis did not focus on truly Placebo consumers).

## Appendix C: Search model

### C.1 Proposition: Logistic uncertainty

The logistic p.d.f. and c.d.f. with unit scale parameter are given by  $f(x) = \frac{e^{-(x-\mu)}}{(1+e^{-(x-\mu)})^2}$  and  $F(x) = \frac{1}{1+e^{-(x-\mu)}}$  respectively. We seek the solution to equation

$$z = -c + \int_z^\infty x dF_j(x) + F_j(z) z$$

**Table 23** Effect of Treatment on *Placebo* Customers' Behavior

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	Bought promoted sku	% Items bought in promotion	% Expenditure on promoted sku's	No. of promoted sku's bought	Expenditure on promoted sku's	No. of non-promoted sku's bought	Expenditure on non-promoted sku's
Treatment	0.012	0.006	0.002	0.054	0.031	0.115	0.075
(a): OLS std. errors	(0.115)	(0.749)	(0.896)	(0.076) <sup>a</sup>	(0.346)	(0.11)	(0.301)
(b) Clustered std. errors	(0.055) <sup>a</sup>	(0.59)	(0.828)	(0.138)	(0.416)	(0.335)	(0.422)
(c): Wild bootstrap	(0.13)	(0.746)	(0.92)	(0.3)	(0.542)	(0.43)	(0.504)
Constant	0.015	0.289 <sup>c</sup>	0.274 <sup>c</sup>	-0.266	-0.688	-0.203	0.042
	(0.616)	(0.000)	(0.000)	(0.352)	(0.744)	(0.524)	(0.938)
Controls & Store Group Fixed Effects	✓	✓	✓	✓	✓	✓	✓
R-Squared (Within)	0.01	0.008	0.009	0.015	0.012	0.026	0.03
N: Observations:	8,336	2,398	2,398	8,336	8,336	8,336	8,336

Notes: <sup>a</sup> $p \leq 0.10$ , <sup>b</sup> $p \leq 0.05$ , <sup>c</sup> $p \leq 0.01$ . p-values are in parentheses. Standard errors for p-values in (b) and (c) are clustered at the store level. p-values for the treatment effect in (c) and the constant are derived from the wild bootstrap procedure with 1,000 replications

with respect to  $z$ . Plugging in the expressions above yields

$$z = -c + \int_z^\infty \frac{x e^{-(x-\mu)}}{(1 + e^{-(x-\mu)})^2} dx + \frac{z}{1 + e^{-(z-\mu)}}. \quad (19)$$

Integration by parts yields

$$\int_z^\infty \frac{x e^{-(x-\mu)}}{(1 + e^{-(x-\mu)})^2} dx = \log(e^z + e^\mu) + \frac{z}{1 + e^{z-\mu}} - z \quad (20)$$

and the reservation value Eq. 19 becomes

$$\begin{aligned} z &= -c + \log(e^z + e^\mu) + \underbrace{\frac{z}{1 + e^{z-\mu}} - z + \frac{z}{1 + e^{-(z-\mu)}}}_{=0} \\ \Leftrightarrow z &= -c + \log(e^z + e^\mu) \\ \Rightarrow z^* &= \log\left(\frac{e^\mu}{e^c - 1}\right) = \mu - \log(e^c - 1) \end{aligned}$$

and the solution is unique for  $\mu, c \in \mathbb{R}$

## C.2 Theorem: Contraction mapping

Let  $u$  be a random variable with continuous p.d.f. and c.d.f.  $f(\cdot)$ ,  $F(\cdot)$  respectively. The indifference condition is given by

$$\begin{aligned} z^* &= -c + Pr(u \geq z^*) E[u | u \geq z^*] + Pr(u < z^*) z^* \\ &= -c + \int_{z^*}^\infty u f(u) du + z^* F(z^*) \end{aligned}$$

Define  $\Gamma(z) = -c + \int_z^\infty u f(u) du + z F(z)$ . Under standard continuity assumptions,  $\Gamma(z)$  is a contraction mapping if  $\Gamma'(z) \in [0, 1)$ . In our case,

$$\begin{aligned} \Gamma'(z) &= \frac{d}{dz} \left( -c + \int_z^\infty u f(u) du + z F(z) \right) \\ &= 0 - z f(z) + z f(z) + F(z) \\ &= F(z) \end{aligned}$$

which is bounded between zero and one. The use of the Leibniz integral rule implies integration must be interchangeable with differentiation. The contraction mapping applies for a large class of differentiable distributions, as long as  $\int_{z_n}^\infty u f(u) du$  is finite  $\forall z_n \in \mathbb{R}$ , which is also implied by the original Weitzman (1979) model. So, the theorem applies to most distributions used in empirical work.

Using the proposition above, it is easy to show that in the case of mixture of logistics given by Eq. 8, the reservation value can be found through contraction

$$\Gamma(z) = -c + \omega_{ijt}^K \log(e^z + e^{v_{ijt} + \gamma^D}) + (1 - \omega_{ijt}^K) \log(e^z + e^{v_{ijt} + \gamma^S})$$

where  $\omega_{ijt}^\kappa$  is consumer  $i$ 's belief associated with finding a deep discount for a given promotional history  $\kappa \in \{S, D\}$ .

### C.3 Likelihood and estimation

We now characterize the likelihood of an alternative being chosen, which involves adding over search sequences. First, we rank the inside alternatives by their reservation values such that  $z_1 > z_2 > \dots > z_n$ , where  $n$  is equal to the number of inside alternatives in the choice set. We depict the potential search paths consistent with a choice of alternative  $j$  in the diagram of Fig. 7, where searching an additional option corresponds to a lateral movement, and a downward one depicts the purchase of alternative  $j$ .<sup>41</sup> For a consumer to be willing to search option  $j$  with reservation value  $z_j$ , she must have inspected options with higher reservation values before and have found that it was worthwhile searching option  $j$  nonetheless. The reason is that options are ordered by their reservation values, and so if a consumer did not search option  $j - 1$  then she prefers not to search option  $j$  either. The sequence of events leading the consumer to arrive to node  $j$  is given by

$$\begin{aligned} z_1 > u_0 \wedge z_2 > \max \{u_0, u_1\} \wedge z_3 > \max \{u_0, u_1, u_2\} \wedge \dots \wedge z_j > \max \{u_0..u_{j-1}\} \\ = z_j > \max \{u_0..u_{j-1}\} \end{aligned} \quad (21)$$

The identity above can be shown by induction. If a consumer searched option 2 for example, then  $z_2 > \max \{u_0, u_1\}$ . This implies the consumer also searched option 1 because

$$z_2 > \max \{u_0, u_1\} \Rightarrow z_1 > u_0$$

since  $z_2 < z_1$ . For the consumer to prefer option  $j$  to the options searched before, we require  $u_j > \max \{u_0..u_{j-1}\}$ , and so a consumer searches alternative  $j$  and considers it the best option up to that stage if and only if

$$\text{Reach Node } j : \min \{z_j, u_j\} > \max \{u_0..u_{j-1}\} \quad (22)$$

Conditional on searching option  $j$  and preferring it up to that stage, many subsequent search paths can lead to a final choice of  $j$ . For example, the consumer may choose alternative  $j$  without searching any further, or do so after searching option  $j + 1$ , options  $j + 1$  and  $j + 2$ , etc. Let  $\text{Buy}_{j|k}$  be each of such *subsequent* paths, where  $j$  is the chosen product and  $k \geq j$  is the last product searched by the consumer. Then, the probability of choosing option  $j$ , which informs our likelihood function, is equal to

$$\begin{aligned} \Pr(\text{Choose } j) &= \Pr \{ \text{Reach Node } j \wedge (\text{Buy}_j | \text{Reach Node } j) \} \\ &= \Pr \left( \min \{z_j, u_j\} > \max \{u_0..u_{j-1}\} \wedge \left( \bigvee_{k=j}^n \text{Buy}_{j|k} \right) \right) \end{aligned} \quad (23)$$

<sup>41</sup>After this manuscript was accepted, we learned about a result by Choi et al. (2018) that substantially simplifies the estimation of search-based demand models.

We now characterize each of the paths, where movements referred to as ‘down’ and ‘right’ are related to the ones in the Fig. 7:

$$\begin{aligned}
 Buy_{j|j} &= \underbrace{u_j > z_{j+1}}_{Path\ Down_j} \\
 Buy_{j|j+1} &= \underbrace{(\sim Path\ Down_j) \wedge u_j > u_{j+1}}_{Path\ Right_j} \wedge \underbrace{u_j > z_{j+2}}_{Path\ Down_{j+1}} \\
 Buy_{j|j+2} &= Path\ Right_j \wedge \underbrace{(\sim Path\ Down_{j+1}) \wedge u_j > u_{j+2}}_{Path\ Right_{j+1}} \wedge \underbrace{u_j > z_{j+3}}_{Path\ Down_{j+2}} \\
 &\vdots \\
 Buy_{j|k} &= \begin{cases} \left( \bigwedge_{l=j}^{k-1} Path\ Right_l \right) \wedge Path\ Down_k, & j \leq k < n \\ \left( \bigwedge_{l=j}^{k-1} Path\ Right_l \right), & j \leq k = n \end{cases}
 \end{aligned}$$

We have characterized the likelihood function. It remains to maximize it with respect to parameters, conditional on the data. Because utilities are probabilistic, we use simulation to generate  $u$ 's and construct the likelihood. Moreover, the need to investigate multiple search paths led us to employ 10,000 draws per choice-alternative.

In order to account for heterogeneity in search sequences, we add a noise parameter  $\eta \sim N(0, 1)$  to the reservation values. For example, in some circumstances, consumers may not include some products in their consideration sets, which is equivalent to those products featuring very low reservation values. This assumption also provides the demand function with smoothness for purposes of the counterfactual analysis.

An additional difficulty with ‘accept/reject choice simulation’ is that small changes in parameters do not affect simulated outcomes, even for large sets of draws.<sup>42</sup> Moreover, the log-likelihood function exhibits saddle points that make finding the global maximum challenging.

We implement a patterned grid search across a wide range of parameter values, and ensure that the bounds set for the parameters were never achieved during the estimation procedure. Calculation of the standard errors required additional smoothing. For this purpose, following McFadden (1989), we smoothed out the likelihood function by use of a kernel function, which in our case is analogous to adding a low-variance extreme-value noise to each  $u$  and  $z$  component.<sup>43</sup> For illustration purposes, suppose we observe option  $n - 1$  being chosen. The probability of this choice is

$$\begin{aligned}
 Pr(Choose_{n-1}) &= Pr \left\{ \min \{z_{n-1}, u_{n-1}\} > \max \{u_0, \dots, u_{n-2}\} \wedge \left( \bigvee_{k=n-1}^n Buy_{n-1|k} \right) \right\} \\
 &= Pr \{ \min \{z_{n-1}, u_{n-1}\} > \max \{u_0, \dots, u_{n-2}\} \\
 &\quad \wedge (u_{n-1} > z_n \vee (u_{n-1} < z_n \wedge u_{n-1} > u_n)) \}
 \end{aligned}$$

<sup>42</sup>See Train (2009) (Sec. 5.6.2) for a careful exposition of this issue.

<sup>43</sup>See Honka and Chintagunta (2017) for an application within the search framework.

Given a parameter guess, we generate  $R$  sets of simulations of  $u$ 's. For each set  $r$ , we calculate

$$p^r(\text{Choose}_{n-1}) = K(\min\{z_{n-1}, u_{n-1}^r\} - \max\{u_0^r, u_{n-2}^r\}) \\ \cdot [K(u_{n-1}^r - z_n) + K(z_n - u_{n-1}^r) \cdot K(u_{n-1}^r - u_n^r)]$$

where

$$K(x) = \frac{1}{1 + \exp(-\frac{x}{\sigma})}$$

is the logistic kernel with smoothing parameter  $\sigma = 0.001$ . We used the smoothing parameter to calculate standard errors. During estimation, we used  $K(x) = 1$  ( $x > 0$ ) instead, because the grid search algorithm does not require smoothing out the objective function.

Finally, we average across simulation results to calculate the choice probability, i.e.

$$Pr(\text{Choose}_{n-1}) \simeq \frac{1}{R} \sum_{r=1}^R p^r(\text{Choose}_{n-1}).$$

McFadden (1989) characterizes the estimator above as well as its consistency in detail.

## References

- Anderson, E.T., & Simester, D.I. (2004). Long-Run Effects of promotion depth on new versus established customers: Three field studies. *Marketing Science*, 23(1), 4–20.
- Becker, G.S., Grossman, M., Murphy, K.M. (1991). Rational addiction and the effect of price on consumption. *American Economic Review*, 81(2), 237–241.
- Becker, G.S., & Murphy, K.M. (1988). A theory of rational addiction. *Journal of Political Economy*, 96(4), 675–700.
- Bertrand, M., Duflo, E., Mullainathan, S. (2004). How much should we trust Differences-In-Differences estimates?. *The Quarterly Journal of Economics*, 119(1), 249–275.
- Blattberg, R.C., Briesch, R., Fox, E.J. (1995). How promotions work. *Marketing Science*, 14(3), G122–G132.
- Bolton, R.N. (1989). The relationship between market characteristics and promotional price elasticities. *Marketing Science*, 8, 153–159.
- Boulding, W., Lee, E., Staelin, R. (1995). Mastering the mix: do advertising, promotion and sales force activities lead to differentiation. *Journal of Marketing Research*, 31(2), 159–172.
- Briesch, R.A., Krishnamurthi, L., Mazumdar, T., Raj, S.P. (1997). A comparative analysis of reference price models. *Journal of Consumer Research*, 24(2), 202–214.
- Bronnenberg, B.J., Kim, J.B., Mela, C.F. (2016). Zooming in on choice: How do consumers search for cameras online?. *Marketing Science*, 35(5), 693–712.
- Cabral, L.M., & Villas-Boas, M. (2005). Bertrand supertraps. *Management Science*, 51(4), 599–613.
- Cameron, A.C., Gelbach, J.B., Miller, D.L. (2008). Bootstrap-based improvements for inference with clustered errors. *Review of Economics and Statistics*, 90(3), 414–427.
- Cameron, A.C., & Miller, D.L. (2015). A practitioner's guide to cluster-robust inference. *Journal of Human Resources*, 50, 317–372.
- Chan, T., Narasimhan, C., Zhang, Q. (2008). Decomposing promotional effects with a dynamic structural model of flexible consumption. *Journal of Marketing Research*, 45(4), 487–498.
- Choi, M., Dai, A.Y., Kim, K. (2018). Consumer search and price competition. *Econometrica*, 86(4), 1257–1281.
- Davis, S.J., Inman, J., McAlister, L. (1992). Promotion has a Negative Effect on Brand Evaluations-Or Does it? Additional Disconfirming Evidence. *Journal of Marketing Research*, 29(1), 143–148.



- Dodson, J.A., Tybout, A.M., Sternthal, B. (1978). Impact of deals and deal retraction on brand switching. *Journal of Marketing Research*, 15(1), 71–81.
- Dube, J.-P., Hitsch, G.J., Rossi, P.E. (2010). State dependence and alternative explanations for consumer inertia. *The RAND Journal of Economics*, 41(3), 417–445.
- Ehrenberg, A.S.C., Hammond, K., Goodhardt, G.J. (1994). The After-Effects of Price-Related consumer promotions. *Journal of Advertising Research*, 34(4), 11–21.
- Erdem, T., Imai, S., Keane, M.P. (2003). Brand and quantity choice dynamics under price uncertainty. *Quantitative Marketing and Economics*, 1(1), 5–64.
- Erdem, T., Keane, M.P., Sun, B. (2008). A dynamic model of brand choice when price and advertising signal product quality. *Marketing Science*, 27(6), 1111–1125.
- Freimer, M., & Horsky, D. (2008). Try it, you will like it - Does consumer learning lead to competitive price promotions?. *Marketing Science*, 27(5), 796–810.
- Gedenk, K., & Neslin, S.A. (1999). The role of retail promotion in determining future brand loyalty: Its effect on purchase event feedback. *Journal of Retailing*, 75(4), 433–459.
- Gedenk, K., Neslin, S.A., Ailawadi, K.L. (2010). *Sales Promotion in: Retailing in the 21st Century: Current and Future Trends*. Berlin: Springer-Verlag Eds.
- Guadagni, P.M., & Little, J.D. (1983). A logit model of brand choice calibrated on scanner data. *Marketing Science*, 2(3), 203–238.
- Hauser, J.R., & Wernerfelt, B. (1990). An evaluation cost model of consideration sets. *Journal of Consumer Research*, 16(4), 393–408.
- Honka, E. (2014). Quantifying search and switching costs in the US auto insurance industry. *The RAND Journal of Economics*, 45(4), 847–884.
- Honka, E., & Chintagunta, P. (2017). Simultaneous or Sequential? Search Strategies in the U.S. Auto Insurance Industry. *Marketing Science*, 36(1), 21–42.
- Imbens, G. (2011). Experimental design for unit and cluster randomized trials, 3ie, International Initiative for Impact Evaluation.
- Imbens, G., & Kolesar, M. (2015). Robust Standard Errors in Small Samples: Some Practical Advice. Working Paper.
- Imbens, G.W., & Rubin, D.B. (2015). *Causal inference in statistics, social, and biomedical sciences*. Cambridge: Cambridge University Press.
- Jedidi, K., Mela, C.F., Gupta, S. (1999). Managing advertising and promotion for long-run profitability. *Marketing Science*, 18(1), 1–22.
- Jones, J.M., & Zufryden, F.S. (1980). Adding explanatory variables to a consumer purchase behavior model: an exploratory study. *Journal of Marketing Research*, 17(3), 47–53.
- Ke, T.T., Shen, Z.-J.M., Villas-Boas, J.M. (2016). Search for information on multiple products. *Management Science*, 62(12), 3576–3603.
- Kehoe, P., & Midrigan, V. (2015). Prices are sticky after all. *Journal of Monetary Economics*, 75, 35–53.
- Kim, J.B., Albuquerque, P., Bronnenberg, B.J. (2010). Online demand under limited consumer search. *Marketing Science*, 29(6), 1001–1023.
- Kopalle, P.K., Mela, C.F., Marsh, L. (1999). The dynamic effect of discounting on sales: Empirical analysis and normative pricing implications. *Marketing Science*, 18(3), 317–332.
- Lal, R., & Matutes, C. (1994). Retail pricing and advertising strategies. *Journal of Business*, 67(3), 345–370.
- McFadden, D. (1989). A method of simulated moments for estimation of discrete response models without numerical integration. *Econometrica*, 57(5), 995–1026.
- Mela, C.F., Gupta, S., Lehmann, D.R. (1997). The long-term impact of promotion and advertising on consumer brand choice. *Journal of Marketing Research*, 34(2), 248–261.
- Narasimhan, C., Neslin, S.A., Sen, S.K. (1996). Promotional elasticities and category characteristics. *Journal of Marketing*, 60(2), 17–30.
- Neslin, S.A., & Shoemaker, R.W. (1989). An alternative explanation for lower repeat rates after promotion purchases. *Journal of Marketing Research*, 26(2), 205–213.
- Neslin, S.A., & Van Heerde, H.J. (2009). Promotion dynamics. *Foundations and Trends in Marketing*, 3(4), 177–268.
- Pollak, R.A. (1970). Habit formation and dynamic demand functions. *Journal of Political Economy*, 78(4, Part 1), 745–763.
- Raju, J.S. (1992). The effect of price promotions on variability in product category sales. *Marketing Science*, 11(3), 207–220.

- Rubin, D.B. (1973). Matching to remove bias in observational studies. *Biometrics*, 29, 159–183.
- Rubin, D.B. (1979). Using multivariate matched sampling and regression adjustment to control bias in observational studies. *Journal of the American Statistical Association*, 74(366), 318–328.
- Scott, C.A. (1976). The effects of trial and incentives on repeat purchase behavior. *Journal of Marketing Research*, 13(1), 263–269.
- Seiler, S. (2013). The impact of search costs on consumer behavior: a dynamic approach. *Quantitative Marketing and Economics*, 11(2), 155–203.
- Seiler, S., & Pinna, F. (2017). Estimating search benefits from Path-Tracking data: Measurement and determinants, Marketing Science, Forthcoming.
- Shoemaker, R.W., & Shoaf, R.F. (1977). Repeat rates of deal purchases. *Journal of Advertising Research*, 17(2), 47–53.
- Shugan, S.M. (1980). The cost of thinking. *Journal of Consumer Research*, 7(2), 99–111.
- Spinnewyn, F. (1981). Rational habit formation. *European Economic Review*, 15(1), 91–109.
- Train, K.E. (2009). *Discrete choice methods with simulation*. Cambridge: Cambridge University Press.
- Tuchman, A., Nair, H., Gardete, P.M. (2017). Complementarities in consumption and the consumer demand for advertising, Working Paper.
- Villas-Boas, S.B., & Villas-Boas, J.M. (2008). Learning, forgetting, and sales. *Management Science*, 54(11), 1951–1960.
- Vuong, Q.H. (1989). Likelihood ratio tests for model selection and non-nested hypotheses. *Econometrica*, 57(2), 307–333.
- Weitzman, M.L. (1979). Optimal search for the best alternative. *Econometrica*, 47(3), 641–654.
- Zenor, M.J., Bronnenberg, B.J., McAlister, L. (1998). The impact of marketing policy on promotional price elasticities and baseline sales. *Journal of Retailing and Consumer Services*, 5(1), 25–32.
- Zou, D. (2014). Intertemporal pricing of new products, Incentivizing Consumer Learning and Inertia, Working Paper.
- Zubizarreta, J.R. (2012). Using mixed integer programming for matching in an observational study of kidney failure after surgery. *Journal of the American Statistical Association*, 107(500), 1360–1371.

## Affiliations

Andrés Elberg<sup>1</sup> · Pedro M. Gardete<sup>2</sup>  · Rosario Macera<sup>3</sup> · Carlos Noton<sup>4</sup>

✉ Pedro M. Gardete  
gardete@stanford.edu

Andrés Elberg  
aelberg@uc.cl

Rosario Macera  
rmacera@uandes.cl

Carlos Noton  
cnoton@dii.uchile.cl

<sup>1</sup> Department of Economics and School of Management, Diego Portales University, Santiago, Chile

<sup>2</sup> Graduate School of Business, Stanford University, Stanford, CA, USA

<sup>3</sup> School of Business and Economics, Universidad de los Andes, Santiago, Chile

<sup>4</sup> Center for Applied Economics, Department of Industrial Engineering, University of Chile, Santiago, Chile